61-102

274063

ELECTRONIC DEVICES

By

M. S. Kaufman and G. M. Yankin

FTD-TT- 61-102/1+2

UNEDITED ROUGH DRAFT TRANSLATION

ELECTRONIC DEVICES (SECOND EDITION)

BY: M. S. Kaufman and G. M. Yankin

English Pages: 653

PART I OF II PARTS (Pages 1-393)

THIS TRANSLATION IS A RENDITION OF THE ORIGINAL FOREIGN TEXT WITHOUT ANY ANALYTICAL OR EDITORIAL COMMENT. STATEMENTS OR THEORIES ADVOCATED OR IMPLIED ARE THOSE OF THE SOURCE AND DO NOT NECESSARILY REFLECT THE POSITION OR OPINION OF THE FOREIGN TECHNOLOGY DIVISION.	PREPARED BY: TRANSLATION SERVICES BRANCH FOREIGN TECHNOLOGY DIVISION WP-AFB, OHIO.

FTD-TT- 61-102/1+2

Date 26 Jan. 1962

S

M. S. KAUFMAN and G. M. YANKIN

ELEKTRONNYYE PRIBORY

Elektronnyye Pribory Gosudarstvennoye Energeticheskoye Izdatel'stvo
Moskva 1960 Leningrad

Pages: 544

FTD-TT-61-102/1+2

The theory of the operation, construction and basic design of electronic and semiconductor devices, as well as the relationship between the parameters and designs of the devices and the most important circuits in which they are applied are set forth in this book. The basic facts of electron theory and electron optics are considered at the start.

The exposition of the material in the book presupposes that the reader is prepared in mathematics, physics, and general technology to the extent provided by the corresponding courses in the electron-tube technical schools.

The book may be used as a text for the course "Electronic Devices" in the electron-tube technical schools, and as an aid for workers concerned with problems of the design, functioning and application of electronic devices.

Kaufman, Mikhail Simonovich, and Yankin, Grigoriy Maksimovich

ELECTRONIC DEVICES

V.I. Shamshur, Editor K.P. Voronin, Technical Editor

Sent to Printer 20 April 1960 Sent to Press 3 August 1960
T-10152 Sheet size 84 x 108; 32-page signatures
27.9 Printer's sheets 31 Printer's billing sheets
First printing, 15,000 copies Price 11 rubles, 80 kopecks,
Price after 1 January 1961, 1 ruble, 18 kopecks
Order Number 2221

Printing House of the State Power Press, Moscow, #10 Shlyuzovaya Wharf.

TABLE OF CONTENTS

FIRST LINE OF TITLE

STOP HERE

STOP HERE

EDITOR'S FOREWORD

The book "Electronic Devices" is herewith published in its second edition as a one-volume work. The very fact that it is being republished indicates its recognition both as a textbook for electron-tube technical schools, and as an aid to workers in the various fields of technology which are connected with the manufacture and application of the devices described in the book.

In the process of re-editing the book for the second edition, we have made use of the experience gained in its use as a textbook for the course in "Electronic Devices" in the electron-tube technical schools, as well as critical observations pertaining to the content, technique of exposition and arrangement of the book and some new material which was published after the appearance of the first edition.

Certain chapters of the book have been shortened considerably at the expense of problems basically concerned with utilization of the devices in specific circuits. On the other hand, the exposition of those questions which pertain to the design features of the devices described in the book and the processes occurring in the latter has been enlarged. The exposition of several sections, including the general problems of electron theory and problems of the operation of certain semiconductor devices and special electron tubes is brought nearer to the contemporary level of the science.

The book does not pretend to give an all-inclusive analysis of the whole complex of physical processes in the types of electronic

devices considered; it limits itself basically to the content provided for students in the electron-tube technical schools. Problems of the design of the devices are set forth with consideration of the level of mathematical preparation of the students. The exposition of these problems has as its object teaching the student the skills required to carry out independent calculations of the separate elements of the electronic devices.

Chapters 1-4, 13, 14, 16 and 17, as well as Sections 15-8 and 15-9 were written by G.M. Yankin; Chapters 5-12 and 15 and Section 4-7 were written by M.S. Kaufman.

R.A. Nilender

Chapter One

FUNDAMENTALS OF ELECTRON THEORY

1-1. STRUCTURE OF THE ATOM. ELECTRONS AND IONS.

According to present-day electronic theory, matter has an atomic structure, i.e., is composed of extremely tiny particles which are called atoms. The atom, in turn, is composed of even smaller electrically charged and neutral particles.

The simplest atomic structure may be found in the hydrogen atom: in the center of the atom is a heavy nucleus with an elementary positive charge – the proton, in which practically the whole mass of the atom is concentrated, while a negatively charged particle, the electron, revolves around the nucleus, under the influence of its attraction. As a result of this and of the coincidence of the center of the positively charged nucleus with the center of the time-average position of the electron, the charges of the nucleus and the electrons cancel each other with the result that the atom as a whole is electrically neutral.

The helium atom has a somewhat more complex structure. Its nucleus contains two protons and two electrically neutral particles which are called neutrons. The mass of the neutron is somewhat greater than that of the proton. Two electrons revolve around the helium nucleus (Fig. 1-2).

The representation of the electron given in Figs. 1-1 and 1-2 is conventional and symbolic. In reality, the electron has no completely definite, rigorously fixed shape, and we cannot ascribe to it the shape of a little ball. But should we want to imagine a little ball with a charge equal to that of the electron and giving rise to the same field as the electron, the radius of such a ball would be $1.4 \cdot 10^{-13}$ cm.

The representations of the orbits shown in these figures are likewise conventional. In this case, when we speak of an orbit, what we really mean is the most probable position of the electron in the atom. In reality, it is impossible to trace the path of the electron around the nucleus.

As the atomic number of the element increases, the complexity of the atomic structure becomes greater and greater. One of the most complex-structured atoms is that of uranium.

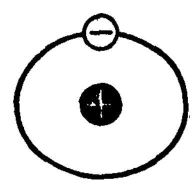

Fig. 1-1. Model of uranium atom.

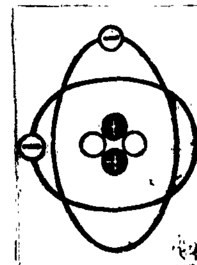

Fig. 1-2. Model of helium atom.

Ninety-two electrons revolve around the positively charged nucleus of the uranium atom at various distances from the center and form so-called "electron shells" around the nucleus. The nucleus itself contains 92 protons and 146 neutrons, so that the atomic weight of uranium is 238.

The number of neutrons in the nucleus of any atom may be determined if we know the atomic weight rounded off to a whole number and the atomic number Z in Mendeleev's periodic chart of the elements:

$$n = A - Z \tag{1-1}$$

The atoms of all the elements are electrically neutral, since the number Z of their protons is equal to that of their electrons.

Electrons may transfer from one electron shell to another, with each such transition accompanied by a change in energy. In

passing to a shell with a smaller radius, the electron gives up part of its energy into the surrounding space as radiant energy, so that its energy becomes smaller. The reverse transition of an electron to shells farther from the nucleus is possible only on application of an external force which, working against the force binding the electron to the nucleus, thereby increases its energy.

If the electron receives sufficient energy, its bond with the nucleus may break. Clearly, the process of complete separation of the electron from the nucleus involves a very great expenditure of energy. For this reason, the internal energy of the atom is greatest upon complete separation of the electron from it. Upon complete separation of the electron, the electrical equilibrium of the atom is disturbed: the atom changes into a positively charged ion, the positive charge of which is equal in magnitude to the negative charge of its lost electron.

The process of complete separation of an electron from an atom or that of the addition of an excess electron to the atom to create either a positive or a negative ion is called the ionization of the atom.

The molecules, atoms or ions of solids are arranged in a definite order and form a so-called crystal space lattice. Four types of lattice may be distinguished: molecular, atomic, ionic and metallic, depending on the particles composing the lattice and the nature of the forces operating between them.

The last two types of lattice are of interest to us.

In ionic lattices, we have a regular, periodically recurring alternation of oppositely charged ions. Electrical forces of attraction act between oppositely charged ions, and electrical forces of repulsion between those which have like charges. The resulting

forces array the ions of both signs in a strictly determined order, by which the structure and density of the crystal lattice are determined. Ionic lattices are typical of the majority of salts. Figure 1-3 shows the elementary cell of the ionic lattice of sodium chloride, which takes the form of a simple cubic lattice.

The lattices typical of metals are most interesting. They are composed of only positively charged ions, among which electrons move freely. The majority of metals are characterized by lattices in which the component particles are extremely close-packed. Certain metals, e.g., tungsten, have lattices which are less tightly packed (Fig. 1-4).

Numerous measurements have made it possible to determine that the charge of the electron is equal to $4.803 \cdot 10^{-10}$ absolute electrostatic units.

Like the proton, the electron has a mass; it is equal to $9.106 \cdot 10^{-28}$ g – almost 2000 times smaller than the mass of the proton. But even with such a small mass, the density of the electron is extremely high, since this mass is enclosed in such a small volume.

The charge of the electron remains constant under all conditions, but its mass remains constant if it is in a state of rest. When an electron moves, its mass changes with changing velocity. The relationship of the mass m of the electron to its velocity v may be expressed by the formula:

$$m = \frac{m_0}{\sqrt{1-\left(\frac{v}{c}\right)^2}}, \qquad (1\text{-}2)$$

where m_0 is the mass of the electron at rest and c is the speed of light. If, in electronic devices, the electron were able to move with a velocity very close to the velocity of light, e.g., with a

velocity of 0.998c, as in the case of radioactivity, it would, according to this formula, have a mass approximately 16 times greater than its mass in the state of rest. In the majority of electronic devices, however, the electrons move with velocities much smaller than that of light, so that the change in the mass of the electron can be disregarded.

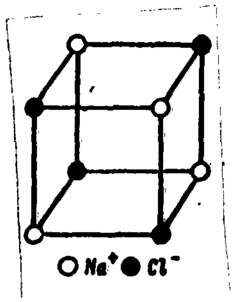

Fig. 1-3. Shape of elementary crystal-lattice cell of sodium chloride.

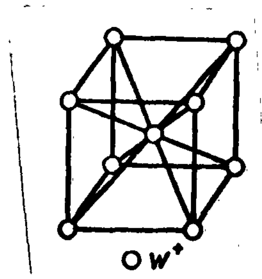

Fig. 1-4. Shape of elementary crystal-lattice cell of tungsten.

The above conceptions of the structure of the atom are, strictly speaking, only approximate, but constitute a useful visual representation. In reality, the physical processes which unfold in the electron shells of the atom are more complex. The so-called quantum theory provides a more exact explanation.

1-2. BASIC CONCEPTS OF THE QUANTUM THEORY.

Planck, one of the creators of the quantum theory, established in 1900 that radiant energy can be emitted and absorbed not uninterruptedly and in any quantity, as implied in the electromagnetic theory of light, but in definite portions (quanta), the magnitude of which can be determined by the equation

$$w = h\nu, \tag{1-3}$$

where ν is the number of oscillations per second (frequency) and h is

a proportionality coefficient (Planck's constant) having the dimensions of action (energy multiplied by time) and therefore known as the quantum of action ($h = 6.62 \cdot 10^{-27}$ erg·sec.). Thus the quanta of energy are multiples of the quantum of action h, which has the same value for all frequencies or wavelengths.

It follows from Equation (1-3) that for low frequencies of oscillation, the energy quanta are also small, i.e., long-wave oscillations possess little energy; at high oscillation frequencies, the energy quanta are large, and the actual processes of radiation unfold with sharp, discontinuous energy changes.

Generalizing the theory of Planck, Bohr established in 1913 that continuous radiation as set forth in the electromagnetic wave theory does not hold true in the case of the atom, and that the laws which apply for phenomena unfolding in the world of large (macroscopic) quantities do not hold true for the atom. In the model of the hydrogen atom constructed by Bohr, the electron can move only in certain completely determined orbits which are permitted by the quantum theory. The electron revolves in such orbits without radiation of energy, with the smallest of these orbits corresponding to the most stable (normal) energy state of the atom, in which it can exist for an indefinite time. The transition of an electron from a given orbit to another, i.e., from a given energy level to another, is accompanied by either absorption or emission of definite quantities of energy, that is, quanta. In transferring to a level situated closer to the nucleus, the electron gives up a part of its energy in the form of radiation to the surrounding space, with the result that its energy becomes smaller. The reverse transition of the electron to a level farther from the nucleus is possible only under the influence of an external force which does work against the forces attracting the electron toward the nucleus and

thereby increases its energy. Consequently, the largest energy is necessary to separate the electron from the atom if the electron is at the lowest energy level. If, however, the electron is situated in an external energy level far from the nucleus, less energy is necessary for its complete separation from the nucleus. The energy which is necessary for complete separation of an electron from an atom is called the energy of ionization.

No more than two electrons may be found simultaneously on each energy level of an atom or system of atoms (Pauli exclusion principle). For this reason, the transition of an electron to a given level is possible only provided that the energy level is free.

Contemporary quantum theory (quantum mechanics), which studies the laws of motion and interaction of particles of atomic dimensions and very small mass (electrons, protons etc.), has established that the most probable situations of the electrons revolving around the nucleus of any atom are given by spheres which correspond to Bohr orbits, but are not the only possible orbits but only the most probable. The electrons in the atom may thus be found at any distance from the nucleus, but they must at the same time occupy only fully defined energy levels; they may not occur at intermediate levels, which are forbidden to them. The presence of forbidden levels has been confirmed experimentally with great certainty, though the cause of the exclusion is not yet clear.

A completely determined amount of energy equal to a whole number of quanta hv corresponds to each energy level. For this reason, the energy difference between any neighboring levels is always equal to hv. Thus the energy of the electrons in the atom can assume not all, but only fully determined, discrete values.

The electron is a material particle (corpuscle). In addition to corpuscular properties, however, it also possesses wave properties at the same time. The corpuscular properties of the electron may be detected in phenomena which occur in spaces whose linear dimensions exceed the size of the electron considerably. If, however, the linear dimensions of the space (of the object of interaction) are comparable to the dimensions of the electron, its wave properties can become apparent.

Wave and corpuscular properties can also be detected in other atomic particles (protons, neutrons, etc.) but wave properties are manifested more strongly in the electron, which has an extremely small mass, than in other atomic particles.

Thus, quantum mechanics has established that, in contrast to the macroscopic bodies, the most important properties of atomic particles are the discreteness of their interactions and the presence of corpuscular and wave properties.

1-3. FREE AND BOUND ELECTRONS.

Under certain conditions, the outer electrons of certain atoms can receive an additional energy from without which is sufficient for complete separation from the nucleus. As a result, the atoms become positive ions and the detached electrons continue their motion among the neutral atoms.

Electrons which move in interatomic space even for a short interval of time are called free electrons. These electrons occupy higher energy levels than the electrons found in the atom.

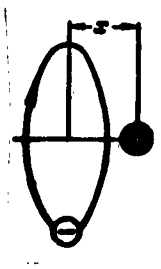

Fig.1-5. Dipole formed when external electric field acts upon an atom.

In the presence of an external electric field, the chaotic movement of the

electrons becomes ordered and directional. An electric current arises as a result of this. It is customary to refer to electrons which are situated within the atom and, unlike free electrons, are closely held to the nucleus, as bound electrons. The bound electrons belong to the atoms and their movements are restricted to the limits of the atoms. The influence of an external electric field on bound electrons is manifested only in a certain displacement of the electrons away from the nucleus and distortion of the electron shell under the influence of the electric forces. Dipoles (Fig. 1-5) form as a result of this, with the positively charged atomic nuclei situated at one end and the electrons on the other. The centers of the nuclei and the time-average positions of the electrons do not coincide in this case, but are displaced by a certain distance x. The greater the strength of the external electric field, the greater will be this shift. Electrons that are bound to the nucleus cannot, of course, create an electric current even in the presence of an external electric field.

1-4. THE VELOCITY AND ENERGY OF ELECTRONS.

Let us consider the changes in the velocity and energy of free electrons which move in a vacuum under the influence of the forces of an electric field.

In general, the force acting on a positive charge q situated in an electric field is directly proportional to the magnitude q of the positive charge and the electric field strength E:

$$F = qE. \qquad (1\text{-}4)$$

If an electric field acts upon a unit positive charge q=1, the equation is simplified:

$$F = E. \qquad (1\text{-}5)$$

Equations (1-4) and (1-5) show that the force F and the

electric field strength E have the same sign and, consequently, the same direction.

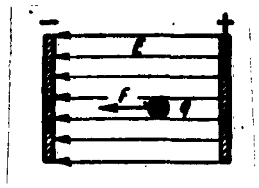

Fig. 1-6. The direction of movement of a positive charge q in an electric field coincides with the direction of the force F and the field vector E.

Fig. 1-7. The movement of an electron in an electric field coincides with the direction of the force F, but is opposite to the direction of the field vector E.

For this reason, the positive charge q will be displaced in the direction of the force F and the force vectors of the electric field E, passing from points of high potential to points of lower potential (Fig. 1-6).

Electric currents are created in electronic devices by the directional movement of free electrons in a vacuum. The force which the electric field exerts upon the free electron is equal in magnitude to the product of the charge on the electron by the electric-field strength, but opposite in direction to the field-strength vector E, since the electron is negatively charged:

$$F = -eE \qquad (1\text{-}6)$$

In contrast to the positive charge, therefore, an electron will travel in the direction of the applied force F under the influence of an electric field, but in the direction opposite to the field vector E, i.e., will travel from a point of lower potential toward a point of higher potential (Fig. 1-7).

The above-described interaction of an electric field with the positive charge or the electron is similar to the familiar

interaction of like and unlike electric charges, in which like charges repel and unlike charges attract.

Let us consider the simplest case of movement of a free electron in an electric field: that in which it traverses a certain potential difference U in a vacuum under the influence of a force $F = -eE$.

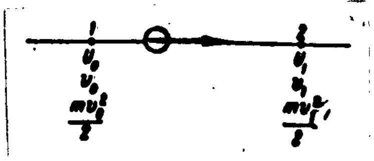

Fig. 1-8. The movement of an electron from a point with lower potential U_0 to a point of higher potential U_1 is accompanied by an increase in its velocity v and kinetic energy $\frac{mv^2}{2}$.

Let us denote the potential of the first point of the field (Fig. 1-8) through which an electron passes with an initial velocity v_0 by U_0 and the potential of the second point to which the electron passes with a velocity of v_1 by U_1.

If the potential U_1 at point 2 is greater than the potential U_0 at point 1, the velocity as well as the kinetic energy of the electron will increase as it passes from point 1 to point 2. The increase in the electron's kinetic energy due to the accelerating potential difference $U_1 - U_0$ through which it has passed between these points of the field may be presented as the difference

$$\Delta w = \frac{mv_1^2}{2} - \frac{mv_0^2}{2}.$$

But on the basis of the law of conservation of energy, the increase in the kinetic energy must be equal to the work done by the forces of the electric field to move the electrons from point 1 to point 2.

This work is equal to the product of the electronic charge by the potential difference $U_1 - U_0$ through which it has passed between points 1 and 2, i.e., $\Delta w = e(U_1 - U_0)$. On this basis, we

can equate the kinetic energy increase of the electron to the work which the electric field has done to displace it:

$$\frac{mv_1^2}{2}-\frac{mv_0^2}{2}=e(U_1-U_0). \qquad (1\text{-}7)$$

If the electron were to move from a point of higher potential to a point of lower potential, it would do work against the field forces at the expense of its initial kinetic energy $mv_0^2/2$ and, after having spent its energy, it would come to rest. On the basis of the same law of conservation of energy, the decrease in the initial kinetic energy of the electron must be equal to the energy increase of the electric field.

Electric fields in which the kinetic energy or velocity of the moving electrons increases are called accelerating fields. However, if the kinetic energy or velocity of the electrons decreases in motion in an electric field, such a field is called a retarding field. The potential differences traversed by the electron are called either accelerating or retarding potential differences accordingly.

It should be remembered, however, that an electron-accelerating field will retard the movement of positively charged particles and vice versa. Moreover, one and the same electric field may be accelerating or retarding, depending upon the direction of movement of the electric charge. For example, if an electron moves against the lines of force of the electric field, the field will be an accelerating field for it; if, however, the electron moves in the direction of the field's force lines, the field will be retarding.

Let us turn to Equation (1-7). If the potential of the first point $U_0=0$ and the initial velocity of the electron at that point is $v_0=0$, then Equation (1-7) will assume the simpler form

$$\frac{mv_1^2}{2} = eU_1. \qquad (1\text{-}8)$$

In this formula, the quantity U_1 is the potential difference between the first and second points of the field, and the quantity $mv_1^2/2$ is the kinetic-energy increase of the electron after traversing the accelerating potential difference U_1 between points 1 and 2.

Since the charge and mass of the electron are constant, it follows from Equations (1-7) and (1-8) that the velocity and kinetic energy of the electron are dependent only on the potential difference U_1 which it traverses. For this reason, the electron's energy is often expressed in units called electron-volts (ev). For example, if an electron traverses a potential difference of 1 v in an accelerating field, it gains an energy of 1 ev.

Solving Equation (1-8) for v_1, we find that

$$v_1 = \sqrt{2\frac{e}{m}U_1}. \qquad (1\text{-}9)$$

Substituting the numerical value of the ratio of the charge of the electron to its mass, $\frac{e}{m} = 5.275 \cdot 10^{17}$ absolute electrostatic units, in Equation (1-9), and expressing the potential difference U_1 in volts (1 v = 1/300 absolute electrostatic unit), we obtain the familiar equation by which the velocity of an electron which passes through a potential difference U_1 may be determined if its initial velocity was zero:

$$v_1 = \sqrt{2 \cdot 5{,}275 \cdot 10^{17} \cdot \frac{1}{300} U_1} \approx 600 \cdot 10^{5} \sqrt{U_1} \text{ [cm/sec]}$$

or, expressing v_1 in km/sec:

$$v_1 \approx 600\sqrt{U_1} \text{ [km/sec]} \qquad (1\text{-}10)$$

In this way, knowing the potential difference U_1 traversed by the electron, we can easily calculate the velocity of the electron by Formula (1-10) if its initial velocity $v_0 = 0$.

1-5. ELECTRONS IN CONDUCTORS, SEMICONDUCTORS AND DIELECTRICS.

As we know, solids are called conductors, semiconductors or dielectrics according to their ability to conduct electric current. This division is somewhat arbitrary, since there is no definite boundary between conductors, semiconductors and dielectrics.

Those substances which conduct practically no electric current at normal temperature even if their atoms have large numbers of dlectrons are called dielectrics. The electrical conductivity of dielectrics usually does not exceed 10^{-10} ohm$^{-1}\cdot$ cm^{-1}. Glass, mica, paraffin, alundum (aluminum oxide Al_2O_3), and a number of other substances are classed as dielectrics.

Metals and their alloys are classed as conductors. The electrical conductivity of conductors is approximately $10^4 - 10^6$ ohm$^{-1}\cdot$cm^{-1}.

Semiconductors take an intermediate position between conductors and dielectrics as regards their electrical conductivity; their electrical conductivities vary from 10^2 to 10^{-10} ohm$^{-1}\cdot$cm^{-1}. Graphite, selenium, silicon and many oxides are typical semiconductors.

In Section 1-2 we considered the possible energy states of an isolated atom. In practice, however, we deal with the energy states of electrons in solids, in which, as a result of interaction of the atoms, the energy levels of the isolated atom are replaced by other, different energy levels--the energy levels of the system of atoms as a whole. Here, as in the case of an isolated atom, the electrons of the solid may occupy only fully determined energy levels in

accordance with quantum theory; they cannot occupy the intermediate energy levels. For this reason, transitions of electrons from certain energy levels to others occur "jumpwise." The number of possible energy levels in a solid is considerably greater than the number of electrons. For this reason, the electrons have greater freedom of movement in an interatomic space than within an isolated atom.

The energy state of a solid is conventionally represented graphically as shown in Fig. 1-9. Energy values expressed in electron-volts are plotted vertically, while the energy levels appear in the form of dashes which occur at definite distances from each other in accordance with the energy differences between the levels. Energy levels that are close to each other combine to form bands or zones.

In this way, the energy state of the electrons is determined by their position in the energy band. Here we are not referring to their geometrical situation in the [solid] body, since the electrons of an energy band are really distributed throughout the volume of the solid.

In explaining the conductivity of bodies, the contemporary electron theory considers three energy bands present in any body: the normal band, which is also called the filled or valence band, the conduction band (band of excitation) and the forbidden band (band of forbidden levels).

The normal band and the conduction band exist in all bodies. Between the normal band and the conduction band lies the forbidden band. The width Δw of the forbidden band (Fig. 1-9) may differ in different bodies; the forbidden band is completely nonexistent in some bodies.

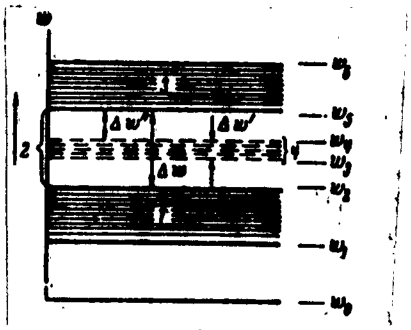

Fig. 1-9. Schematic representation of the energy levels of a body.

The above energy bands make it easy to explain the electrical properties of solids and their classification as conductors, semi-conductors, and dielectrics. Actually, a body will conduct an electric current if its electrons can be displaced under the influence of an external electric field, i.e., can undergo energy changes. This is possible in the presence of free energy levels near occupied levels. If the forbidden band is completely absent from a given body, then the normal band is immediately adjacent to the conduction band (Fig. 1-10a) or may even partially overlap it (Fig. 1-10b) and the valance electrons of the normal band appear in

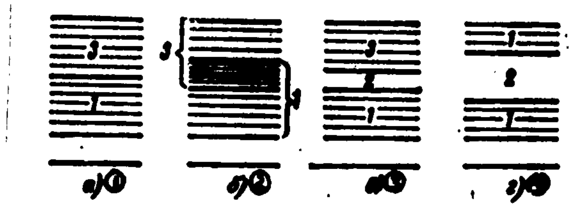

Fig. 1-10. Relative arrangement of energy bands. 1) a = a; 2) б = b; 3) b = c; 4) г = d.

the conduction band, i.e., may participate in the process of conduction. In this case the body is a conductor of electric current.

When all three energy bands exist in a body, the electrical properties of the body will depend upon the width of the forbidden band, which determines the freedom of movement (possibility of energy changes) of the electrons. If the width of the forbidden band is small (Fig. 1-10c), then external disturbances can easily transfer the electrons from the normal band to the conduction band. The greater the external disturbance, the greater the number of electrons that will pass to the conduction band. Such a body is a semiconductor; its conductivity is unstable, since it varies widely with changes in external conditions. For example, an increase in temperature always causes a sharp increase in the conductivity of a semiconductor.

If, however, the width of the forbidden band is great (Fig. 1-10d), the electrons cannot move from the normal band to the conduction band, even under very strong external disturbances. Such a body is incapable of conducting an electric current and is called a dielectric.

The width of the forbidden band in dielectrics is very great and in some instances reaches 15 ev. For this reason, the freedom of movement of electrons in dielectrics is very limited. Because of this there are no electrons in the conduction band of a dielectric at absolute zero; a very small number of the valence electrons of the dielectric acquire increased energies at room temperature because of the thermal agitation of the ions of the crystal lattice and pass from the normal band to the conduction band. But this does not cause a substantial increase in the conductivity of the dielectric.

The width of the forbidden band is much smaller in semiconductors than in dielectrics. Thus, for example, it is 1.11 ev in

silicon and 0.72 ev in germanium. For this reason, the conductivity of semiconductors is considerably greater than that of dielectrics.

The width of the forbidden band can be made smaller artificially by introducing a small amount of another material (an impurity) into the body. In this case, the energy levels of the impurity may fall between the normal band and the conduction band, forming an additional band of width $\Delta w'$ in the band of forbidden levels (Fig. 1-9), and the actual width Δw of the forbidden band becomes smaller, since now the electrons will transfer to the conduction band not from the normal band, but from the band created by the impurity. For this reason, the actual width $\Delta w''$ of the forbidden band (Fig. 1-9) decreases and may have a value of the order of 0.05 ev. Injection of impurities may markedly increase even the conductivity of a dielectric.

Thus the distinction between conductors, semiconductors, and dielectrics is determined not by the fact that the electrons are attached to the atoms more loosely in the conductor than in the semiconductor or the dielectric, but by the fact that the possibility of changes in their energy is very limited in the semiconductor and the dielectric.

1-6. VELOCITY AND ENERGY DISTRIBUTIONS OF ELECTRONS IN CONDUCTORS.

In the absence of an external electric field, free electrons in conductors move in the spaces between the ions of the crystal lattice, in all possible directions and with different velocities, like the molecules of a gas. However, the properties of an ordinary gas and those of the electron gas filling the conductor differ. Actually, the molecules and atoms of an ordinary gas are electrically neutral and their chaotic movements occur in free space. The

electron gas differs from the ordinary gas further in that the electrons are not neutral but negative electric charges; they move not in free space but in a space filled with the positive ions which constitute the crystal lattice of the conductor. For this reason the nature of the movements of the electrons in a conductor differ essentially from the nature of the movements of the molecules and atoms of an ordinary gas.

Contemporary electron theory has established that the velocity distribution of the free electrons in a conductor is subject not to the law of molecular distribution of an ordinary gas, but to the laws of quantum statistics, according to which the electrons in the conductor have fully determined velocities even at absolute zero. Such a distribution is represented by curve 1 in Fig. 1-11, from which it follows that the greatest number of free electrons of a conductor have the highest velocity v_1 at absolute zero. At the temperature of absolute zero, a conductor has no electrons whose velocities exceed the velocity v_1. The number of electrons having velocities lower than v_1, however, diminishes with declining velocity. When a conductor is heated, the velocity distribution of the electrons is disturbed: the number of electrons whose velocities are equal to or somewhat less than the maximum velocity v_1 decreases (region abc), and instead of these electrons, others whose velocities are somewhat greater than the velocity v_1 (region dce) appear, indicating only the velocity redistribution of a few of the electrons; the velocities of the major part of the electrons, however, do not change (segment Oa).

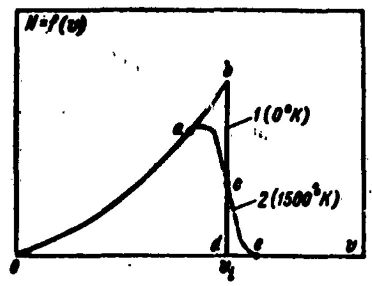

Fig. 11-1. Curves of velocity distribution of the free electrons in conductor at 0 and 1500° K.

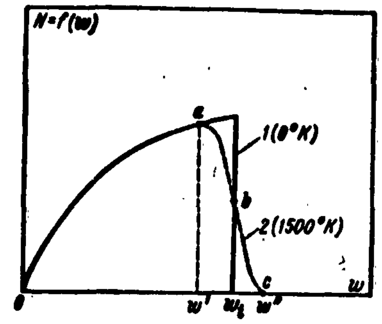

Fig. 1-12. Curves of energy distribution of free electrons in conductor at 0 and 1500° K.

The general appearance of the velocity-distribution curve of the electrons at T = 1500° K in the segment about the velocity v_1 is indicated in Fig. 1-11 by a broken line (curve 2).

The energy distribution of free electrons in a conductor is similar to the velocity distribution described above (Fig. 1-12, curves 1 and 2). The greatest number of a conductor's electrons have the highest energies w_1 at the temperature of absolute zero. The rest of the electrons have energies smaller than the energy w_1. Heating the conductor will, as in the case of the velocity distribution of the electrons, bring about only an energy redistribution of a small number of the electrons: on the one hand, a small number of electrons whose energies are slightly greater than the energy W_1 will appear, and on the other, the number of electrons whose energies are equal to or somewhat smaller than the energy W_1 will diminish.

The shape of the energy distribution curve of the electrons of a heated conductor is indicated by a broken line in segment abc in Fig. 1-12. The energy distribution of electrons in the interval from 0 to W' remains unchanged for both the cold and the heated conductor. For this reason, the curves of distribution coincide in segment O_a.

For this reason, there are more electrons having the greatest energy W_1 in the cold than in the heated conductor. The heated conductor, however, has a few electrons with energies exceeding the energy W_1 (from W_1 to W''). At the temperature of absolute zero, there are no electrons in the conductor with such energies.

The values of the greatest energy W_1 differ in different conductors, and are thus characteristic quantities for them.

1-7. THE ELECTRON WORK FUNCTION.

We ascertained in Section 1-6 that the greatest number of free electrons in the conductor have the greatest energy W_1 at the temperature of absolute zero. Electrons possessing such energies do not, however, leave the conductor. Practically no passage of electrons from the conductor to a vacuum can be observed even at room temperature (293° K), at which a certain number of electrons have energies slightly greater than the energy W_1. This is accounted for primarily by the fact that when some of the electrons leave the conductor, they lose their energies completely and accumulate on the surface of the conductor, forming with the positively charged conductor an electric double layer whose field is directed from the conductor toward the layer of electrons (Fig. 1-13). Such a field is a retarding field for electrons passing out of the conductor; it acts upon them with a force F_1 directed to the conductor.

Theoretical calculations, which have been confirmed by experiment, show that the potential gradient in the electric double layer must be equal to $U = W_1/e$. Electrons having an energy of W_1 in the conductor cannot leave the conductor because this energy is completely lost in crossing the retarding potential difference of the electric double layer ($W_1 = eU$). For this reason, only those electrons whose energy is sufficient to overcome the retarding

action of this field can leave the conductor.

In the second place, the electrons which pass from the conductor to the vacuum induce charges of the opposite sign in the conductor; an electric force of mutual attraction F_2 (force of electric reflection) acts between the electrons and the induced positive charge; it is directed toward the conductor and consequently hinders further movement of the electrons from the conductor into the vacuum (Fig. 1-14).

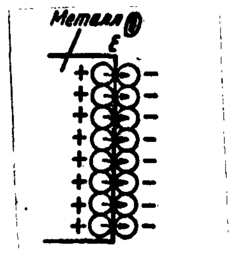

Fig. 1-13. Electric double layer on surface of conductor.
1) Metal.

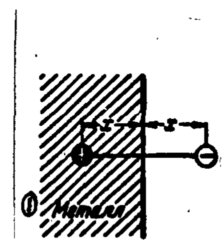

Fig. 1-14. An electron that has left the conductor induces in the conductor a charge of the opposite sign.
1) Metal.

In order to overcome these forces, the electrons must possess energies greater than the greatest energy they may possess in the conductor at a temperature close to absolute zero.

Let us assume for purposes of simplification that a certain resultant force acts upon the electron, tending to hold it back in the conductor. The electron must obviously do a certain work in order to overcome this force and leave the conductor. Since an electrostatic force is acting in this case, the recession of an electron that has left the surface of the conductor to a corresponding distance from it may be compared to the passage of an electron through a certain retarding potential difference between the point of exit from the conductor and the point where the force is prac-

tically equal to zero. Let us denote this potential difference by φ_0. Then the work W_0 which the electron carries out in passing 1
through the potential difference φ_0 may be expressed as the product 2
of the electronic charge and this potential difference: $W_0 = e\varphi_0$. 3

An electron possessing an inherent greatest energy W_1 in the conductor can obviously leave the conductor only provided that an additional energy no smaller than $e\varphi_0$ is transmitted to it from with- 4
out. This smallest additional energy plus the greatest energy W_1 which the electron may possess inside the conductor is the full minimum electron energy W_a for departure from the conductor:

$$W_a = W_i + W_0 = W_i + e\varphi_0.$$

The smallest energy $e\varphi_0$ which the electron must receive from 5
without in order to do work against the forces holding it in the conductor and pass outside the conductor may be determined from the above equality:

$$e\varphi_0 = W_a - W_1. \quad 6$$

The smallest additional energy $e\varphi_0$ that must be expended in 7
order the remove a single electron from a conductor into a vacuum is called the "effective electron work function" or simply "work function," while the full minimum energy W_a of an electron is the full total work function, which indicates that if an electron's energy in the conductor were zero at absolute zero, it would be necessary to transmit to this electron not only the energy $e\varphi_0$ but also the energy W_1 in order to discharge it from the conductor.

Thus, an electron crosses a retarding potential difference in passing from a conductor to a vacuum, with the result that its kinetic energy diminishes. Conversely, in passing from a vacuum into a conductor, it passes an accelerating potential difference and

its kinetic energy increases.

Since an electron passes a retarding potential difference when leaving a conductor, it is customary to say that a more or less sharp potential change occurs at the boundary between the conductor and the vacuum, presenting a barrier or a kind of (potential) threshold, whose height corresponds to the total work function W_a of the electron, to the electron leaving the conductor.

The curve of potential variation close to the surface of a conductor is presented in Fig. 1-15. The potential φ is plotted against the Y-axis, which coincides with the surface of the conductor, while the X-axis is perpendicular to surface of the conductor. The curve shows that the potential undergoes the sharpest change (jump) in the immediate vicinity of the conductor surface.

The passage of electrons out of a conductor is similar to the evaporation of a liquid and the electron work function to the latent heat of evaporation of the molecules of a liquid; here, like the process of molecular evaporation from a liquid, the process of electron "evaporation" from a conductor is accompanied by cooling of the conductor as a result of the carrying away of energy by the "vaporized" electrons.

The value of the work function is different in different conductors; it is a characteristic quantity for each conductor (Table 1-1).

However, the work function is constant even for a given conductor only provided that the surface of the conductor is completely clean. If even a small amount of contamination is introduced, the work function changes sharply. Thus it has been established, for example, that if a sufficiently thin layer of an alkali- or alkaline-earth metal (Cs, Th, Ba, etc.) is applied to the surface

of a conductor, the work function drops noticeably. The reason for the decline in the work function is that atoms of, for example, cesium that are situated on the surface can easily donate their valence electrons to the conductor. The result is that a layer of positive cesium ions forms on the surface of the conductor, contributing to a decline in the work function (Fig. 1-16a).

TABLE 1-1.

Металл	Работа выхода, эв	Металл	Работа выхода, эв
Платина	5,32	Молибден	4,16
Железо	4,72	Тантал	4,10
Никель	4,60	Магний	3,60
Вольфрам	4,60	Торий	3,35
Ртуть	4,52	Кальций	2,71
Углерод	4,38	Стронций	2,60
Медь	4,26	Барий	2,52
Алюминий	4,23	Цезий	1,81

1) Metal; 2) Work function, ev; 3) Platinum; 4) Iron; 5) Nickel; 6) Tungsten; 7) Mercury; 8) Carbon; 9) Copper; 10) Aluminum; 11) Molybdenum; 12) Tantalum; 13) Magnesium; 14) Thorium; 15) Calcium; 16) Strontium; 17) Barium; 18) Cesium.

The atoms which change into positive or negative ions on the surface of the conductor are called electropositive or electronegative relative to the conductor, respectively. The presence of electronegative atoms on the surface of the conductor inhibits the exit of electrons and, consequently, increases the work function (Fig. 1-16b).

The greatest decline in the work function is achieved when sufficiently thin monatomic layers are applied to the surface of the conductor.

The work function increases as the thickness of the layer increases and with sufficiently thick layers it reaches its highest value, which corresponds to the work function of the metal of which the layer is composed. For example, the work function of

pure tungsten is 4.60 ev and that of pure thorium is 3.35 ev. When a monatomic layer of thorium is applied to the surface of

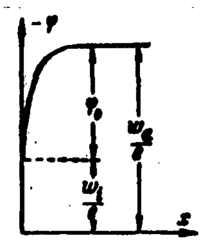

Fig. 1-15. Potential threshold at boundary between conductor and vacuum.

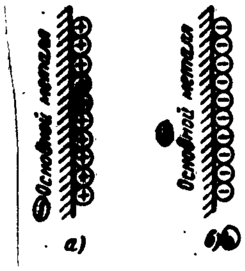

Fig. 1-16. Electropositive and electronegative atoms form positively (a) and negatively (b) charged layers on the surface of a conductor.
1) Base metal.

tungsten, the work function drops to 2.63 ev. If the thickness of the thorium layer is increased, however, the work function begins to rise and reaches, at a certain thickness, the value of 3.35 ev which corresponds to pure thorium. The work function remains unchanged with further increases in the thickness of the thorium layer.

1-8. CONTACT POTENTIAL DIFFERENCE.

Let us consider the processes occurring at the contact boundary between two different conductors. Let conductors A and B have a common surface of contact ab (Fig. 1-17) and free (not contacting) surfaces cd (on conductor A) and ek (on conductor B). A certain number of free electrons moving in the conductors A and B with velocities directed toward the contact surface between the conductors may, in the presence of electrical contact between the conductors, pass through the contact surface from conductor A to conductor B as well as from conductor B to conductor A. If the work functions of the conductors are the same, the number of electrons

passing through the surface from conductor A to conductor B will be equal to the number of electrons passing from conductor B to conductor A. If the work function φ_A of conductor A is smaller than the work function φ_B of conductor B, the free electrons of conductor A will pass to conductor B in greater numbers. Only those electrons which happen to possess the greatest velocities will pass in the opposite direction, from conductor B to conductor A. For this reason, electrons will pass through the contact surface predominantly in one direction, i.e., from conductor A to conductor B.

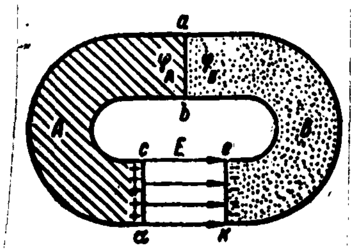

Fig. 1-17. Illustrating definition of contact potential difference.

When an electron leaves conductor A, it must cross a certain retarding potential difference φ_A, i.e., must do a work $e\varphi_A$ at the expense of its kinetic energy. The electron gives this energy up to the electric field of conductor A. As a result, conductor A will become positively charged and its potential will increase. Upon entering conductor B, this electron, which is now moving in an accelerating field, will cross a certain potential difference φ_B greater than φ_A and its kinetic energy will increase at the expense of work done by the electric field of conductor B in displacing it. As a result, conductor B will become negatively charged and its potential will decline. Obviously, the greater the number of free electrons passing through the surface of contact from conductor A to conductor B, the higher will the potential of conductor A become and the lower will be the potential of conductor B, while a certain potential difference will be established between the free surfaces cd and ek. When a state of electrical equilibrium, in which as many electrons pass from conductor A to conductor B as pass in the opposite direction, has been

reached, the electric current ceases to flow through the contact surface of the conductors and the potentials of conductors A and B reach the highest and lowest values, respectively.

Here, as a result of the redistribution of energy between the electric fields of conductors A and B, a potential difference φ_{AB} is created between their free surfaces. Since an electron gives up a smaller energy $e\varphi_A$ to the field as it leaves conductor A and receives a larger energy $e\varphi_B$ from the field as it enters conductor B, the result is that it receives the energy difference

$$e\varphi_B - e\varphi_A = e\varphi_{AB} \tag{1-11}$$

from the field.

The potential difference φ_{AB}, which reaches its maximum value at equilibrium, can be determined directly from Equation (1-11):

$$\varphi_{AB} = \varphi_B - \varphi_A. \tag{1-12}$$

The quantity φ_{AB}, which is also denoted by φ_k* or U_k**, is known as the contact potential difference arising between the free surfaces of two different conductors upon contact. Since the quantities φ_A and φ_B are the work functions of the conductors expressed in equivalent volts, Equation (1-12) gives a direct relationship between the work functions of the conductors and the contact potential difference, indicating that the contact potential difference is equal to the difference between the work functions of the contacting conductors.

---

* [$\varphi_k = \varphi_k = \varphi_{\text{kontaktnaya}} = \varphi_{\text{contact}}$].

** [$U_k = U_k = U_{\text{kontaktnaya}} = U_{\text{contact}}$].

Since the total contact potential difference is established on arrival at electrical equilibrium, at which the electric current flowing through the contact surface declines to zero, it follows that if the free surfaces are joined by a new conductor, the electrical equilibrium will not be disturbed and no current will 2
flow through the connecting conductor, despite the fact that there is a potential difference between the free surfaces. For this reason, the contact potential difference cannot be measured with an ordinary voltmeter, which works when weak currents pass through it. Thus the contact potential difference is not an emf and consequently cannot be used to generate a current, although the electric field between the free surfaces of the metals will act with a 3
force $F = qE$ upon a charge q introduced into it from without and cause it to move in the direction of the forces, which depends on the sign of the charge q.

In many electron tubes, the contact potential difference between the electrodes does not exceed 0.1 v. In some cases, however, it may reach 3.5 v.

It would be possible to take the influence of the contact potential difference into account in designing electronic devices if its magnitude did not change with time. Unfortunately, such allowances are not always possible, since the quantity φ_k varies not only in the various production operations involved in manufacturing the tubes, but also during actual operation of the tubes, when the electrode work function diminishes progressively under the incidence of active atoms from the cathode and brings about a corresponding change in the contact potential difference.

Chapter Two

THE ELEMENTS OF ELECTRON OPTICS

2-1. THE MOTION OF AN ELECTRON IN A HOMOGENEOUS ELECTRIC FIELD.

An electron in an electric field, like any electric charge, is acted upon by a force that causes a change in its motion. The simplest case of this is the motion of an electron in a homogeneous electric field — i.e., a field whose intensity is of the same magnitude and direction at every point. Such a field can be set up by two parallel planes (of sufficiently great extent so that the curvature of the field at the edges may be neglected), between which there is a potential difference.

Let us suppose that an electron enters a homogeneous field of intensity E, and that the electron has an initial velocity v_0, whose direction makes an arbitrary angle with the direction of the field. Let us select a rectilinear coordinate system with its origin at the point where the electron enters the field. Let us choose the positive direction of the Y axis opposite to the direction of the field, and let us direct the X axis perpendicular to the field as shown in Fig. 2-1.

The initial velocity of the electron may be resolved into two components directed along the X axis and the Y axis. Let us denote these components by v_{0x} and v_{0y} respectively.

The electron will be acted upon by a force F of constant magnitude in the positive direction of the Y axis:

$$F = eE. \qquad (2\text{-}1)$$

By the laws of mechanics, a force of constant magnitude and direction causes a uniformly accelerated motion whose acceleration is directly proportional to the magnitude of the force and inversely

proportional to the mass of the body to which the force is applied. Consequently, the acceleration of the electron in the direction of the Y axis may be expressed by the equation

$$a_y = \frac{F}{m} = \frac{eE}{m}. \qquad (2\text{-}2)$$

Since there is no change of potential in the direction of the X axis, there is no force acting on the electron in this direction, and the motion of the electron in this direction is uniformly rectilinear with a velocity equal to v_{0x}. If we consider the motion of the electron in both directions independently, we can determine the path that has been traversed in a time t. The y coordinate of the path of the electron is determined by the formula for the path traversed by a uniformly accelerated body with an initial velocity of v_{0y}:

$$y = v_{0y}t + \frac{a_y t^2}{2}. \qquad (2\text{-}3)$$

If we substitute the value of the acceleration obtained from (2-2) into this expression, we obtain:

$$y = v_{0y}t + \frac{eE}{2m}t^2. \qquad (2\text{-}4)$$

The x coordinate of the path of the electron may be determined from the formula for the path traversed by a body in uniform rectilinear motion with a velocity of v_{0x}:

$$x = v_{0x}t. \qquad (2\text{-}5)$$

Thus, the electron performs two motions simultaneously: a uniformly rectilinear motion in the direction of the X axis and a uniformly variable one in the direction of the Y axis. The orbital equation of the resultant motion may be found by expressing t as a function of x using (2-5) – i.e., $t = x/v_{0x}$ – and substituting the

expression for t in Eq. (2-4):

$$y = \frac{v_{0y}}{v_{0x}} x + \frac{eE}{2mv_{0x}^2} x^2. \qquad (2\text{-}6)$$

Expression (2-6) is the equation of a parabola of the type $y = px^2 + qx$. Consequently, the path of the motion of an electron that enters a homogeneous electric field with an initial velocity of v_0 is a parabola. An example of such a path is shown in Fig. 2-1.

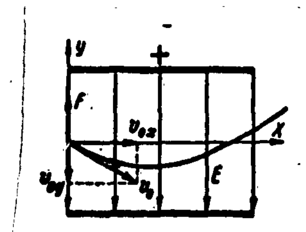

Fig. 2-1. Motion of an electron in a homogeneous electric field (initial velocity of electron directed at an angle α to the lines of force of the field).

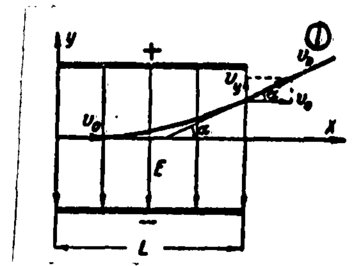

Fig. 2-2. Motion of an electron in a homogeneous electric field (initial velocity of electron directed perpendicularly with respect to the lines of force of the field).
1) v_r.

When it leaves the field, the electron will move rectilinearly along the tangent to the parabola at the point where the field ceased to act on the electron.

In practice, one often has to deal with special cases of the problem discussed above. For example:

1. The electron enters the electric field with an initial velocity that is perpendicular to the direction of the field. This case may be expressed by the relationships $v_{0x} = v_0$ and $v_{0y} = 0$. If we substitute these values for the components of the initial velocity into Eq. (2-6), we obtain:

$$y = \frac{eE}{2mv_0^2} x^2. \qquad (2\text{-}7)$$

In this case, the path of the electron is also a parabola, an

example of which is shown in Fig. 2-2. If we know the initial velocity v_0 of the electron, the field intensity E, and the length of the path in the direction of the X axis L, we can determine at what angle with respect to the X axis the electron will leave the field and the magnitude of the velocity of the electron when it has left the field.

Since the electron is uniformly accelerated in the direction of the Y axis, we can find the velocity of the electron in this direction at any point of the path by means of a formula known from mechanics: $v = \sqrt{2aS}$, where S is the distance traversed by the body in the direction of the accelerated motion. In the case that we are considering, the distance traversed by the electron in the direction of the Y axis is equal to the Y coordinate of the electron at x = L. Consequently, the velocity of the electron in the direction of the Y axis at the moment that it leaves the field is equal to:

$$v_y = \sqrt{2ay}, \tag{2-8}$$

where, according to Eq. (2-7), $y = \frac{eE}{2mv_0^2}L^2$.

If we substitute the value of the acceleration obtained from Eq. (2-2), we have:

$$v_y = \sqrt{2\frac{eE}{m}\frac{eE}{2mv_0^2}L^2} = \frac{eEL}{mv_0}. \tag{2-9}$$

The velocity of the electron in the direction of the X axis does not change (remains equal to v_0). Thus, since we know the magnitude of the components of the velocity of the electron at the point where it leaves the field, we can determine the direction and magnitude of the resultant velocity:

$$\operatorname{tg}\alpha = \frac{v_y}{v_0} = \frac{eEL}{mv_0^2};\ * \qquad (2\text{-}10)$$

$$v_p = \sqrt{v_y^2 + v_0^2}.\ ** \qquad (2\text{-}11)$$

Consequently, an electron that enters a homogeneous electric field at an initial velocity perpendicular to the field intensity, moving along a parabola, changes both the direction and magnitude of its velocity.

2. The electron enters a homogeneous field at an initial velocity directed along the lines of force of the field: $v_{0y} = v_0$; $v_{0x} = 0$. Substituting the relationships into Eqs. (2-4) and (2-5), we obtain the equations for the motion of the electron in the field according to the coordinate system that we selected:

$$y = v_0 t + \frac{eE}{2m} t^2;\ x = 0. \qquad (2\text{-}12)$$

As is evident from these equations, the path of the motion of the electron is a straight line coinciding with the Y axis (there is no displacement of the electron in the direction of the X axis because there are no forces acting in this direction).

If the initial velocity of the electron is directed oppositely to the direction of the field and, consequently, coincides with the direction of the force acting on the electron, the electron will be uniformly accelerated by the action of this force. But if the direction of the initial velocity coincides with the direction of the field, the electron will be acted upon by a force in a direction opposite to the direction of its motion; and, consequently, the electron will be decelerated. Therefore, after it has traveled a certain distance, the velocity of the electron will be

---

* [tg = tan].

** [$v_p = v_r = v_{\text{resul'tiruyushchiy}} = v_{\text{resultant}}$].

zero. This motion is described by the equations for uniformly decelerated motion:

$$y = \frac{eE}{2m} t^2 - v_0 t; \tag{2-13}$$

$$v = v_0 - \frac{eE}{m} t. \tag{2-14}$$

3. At the moment that the forces of the electric field begin to act on the electron, its initial velocity is equal to zero — i.e., $v_{0x} = 0$ and $v_{0y} = 0$. Then Eqs. (2-4) and (2-5) will read as follows:

$$y = \frac{eE}{2m} t^2; \; x = 0. \tag{2-15}$$

These equations describe a uniformly accelerated motion in the direction of the Y axis when the initial velocity is equal to zero.

2-2. THE MOTION OF AN ELECTRON IN NONHOMOGENEOUS ELECTRIC FIELDS. ELECTROSTATIC ELECTRON LENSES.

A homogeneous electric field, strictly speaking, can be created between flat, infinitely large plates, the lines of force between which can be considered straight, perpendicular to the plates and the equipotential surfaces, and as having the same density at all points of the field (Fig. 2-3). But in practice, one must deal with nonhomogeneous fields. Nonhomogeneous fields are created between various kinds of curved surfaces, the lines of force between which are curved and have different density at different points of the field. The equipotential surfaces of a nonhomogeneous field are curved; however, the lines of force are directed along the normal to the equipotential surfaces at any point of the field (Fig. 2-4). In a nonhomogeneous electric field, the field intensity E and the force F acting on the electron change both in magnitude

and direction when the electron moves from point to point.

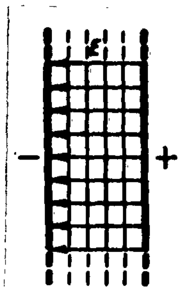

Fig. 2-3. The lines of force (indicated by arrows) and the equipotential surfaces (thin lines) of the homogeneous electric field between parallel plates.

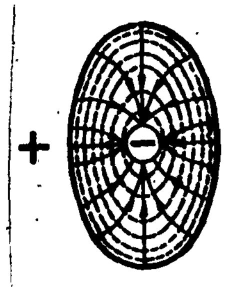

Fig. 2-4. The lines of force (indicated by arrows) and the equipotential surfaces (broken lines) of a nonhomogeneous electric field.

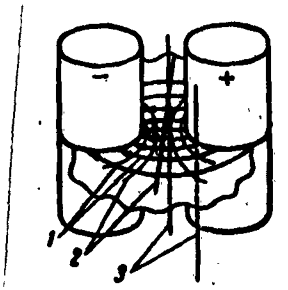

Fig. 2-5. Illustration of the concept of a nonhomogeneous plane field.
1) Lines of force; 2) equipotential surfaces; 3) generators.

The simplest and most easy to analyze case of a nonhomogeneous electric field is the so-called plane field, in which the equipotential surfaces themselves are curved, but their generators are mutually parallel straight lines (Fig. 2-5). The motion of electrons in nonhomogeneous electric fields is generally more complicated than the motion in homogeneous fields; however, the motion of electrons in plane nonhomogeneous fields may be described comparatively easily. The difficulties that arise in the calculation of the path of the motion of the electron occur because the electron, which has an inertial mass, acquires an initial velocity owing to the forces acting upon it; and the direction of this velocity diverges from that of the lines of force as the electron moves through the field. As a result, the calculation of the path of the motion of the electron becomes difficult.

In the case of plane nonhomogeneous electric fields, an electron whose initial velocity $v_0 = 0$ will also not move along the lines of force of the field; but, in spite of this, its motion will

be confined to one plane perpendicular to the generators of the equipotential surfaces; the electron cannot move from one plane to another during the course of its motion since the intensity and the potential of the electric field do not change in the direction of the generators; consequently, no force acts upon the electron in this direction.

The simplest case of a nonhomogeneous plane field is the electric field between the surfaces of sufficiently long coaxial (with coincident axes) cylinders (Fig. 2-6). In this case, the equipotential surfaces are also cylindrical, and the lines of force are directed along the radii perpendicularly to the surfaces of the cylinders. Therefore, the density of the lines of force changes in the radial direction (in this direction, the field is nonhomogeneous) and remains constant in the direction of the axis or any generator of the cylindrical equipotential surfaces (in this direction, the field is homogeneous). In such plane fields, the nature of the potential distribution does not vary when passing from one cross section to another. Electrons in this type of field will move in the radial direction in an accelerating or decelerating field, depending on the direction of the lines of force of the field and the initial velocity of the electrons.

Electrostatic electron lenses. By selecting the magnitude and direction of the initial velocity of the electrons, and also the magnitude and direction of the intensity of the electric field, the direction of the motion of the electrons may be varied at will — i.e., the electrons can be made to move along a previously calculated path, just as a light ray can be made to change its direction by selection of its initial direction and the corresponding optical

media.

A sufficiently narrow stream of electrons moving in one direction is called an electron beam. There is a close analogy between electron beams and light rays, since the motion of the electrons of a beam in an electrostatic field is similar to the propagation of light in a medium with a continuously varying refractive index. Electron beams, like light rays, can be reflected, refracted, and focused. This is explained by the fact that the laws that hold for optical lenses are also applicable to electrostatic lenses, which are electric fields of definite configuration that are created by a system of electrodes with various potentials.

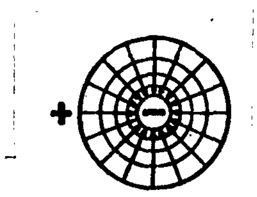

Fig. 2-6. Nonhomogeneous field between coaxial cylinders.

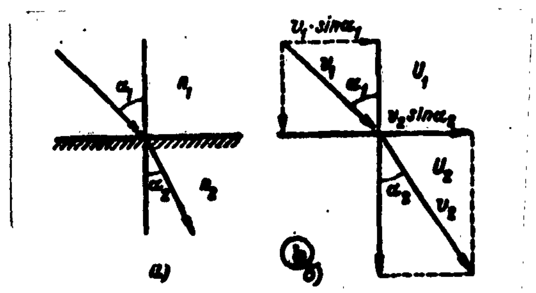

Fig. 2-7. Analogy between the law of refraction for a light ray (a) and an electron beam (b).

One of the fundamental laws of geometrical optics is the law of the refraction of a light ray which passes from one medium with a refractive index of n_1 into another medium with a refractive index of n_2 (Fig. 2-7a):

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2, \qquad (2\text{-}16)$$

where α_1 and α_2 are the angles of incidence and refraction of the light ray.

This law is also applicable to the electron beam. In fact, if

an electron beam passes through a region of space at a potential of U_1 and then enters a region of space at a potential of U_2, the electron will undergo refraction at the interface of the regions. This refraction is caused by the fact that, when the electron passes from the region at the lower potential into the region at the higher potential, the component of the velocity perpendicular to the interface of the regions increases; while the component directed along the normal to the interface remains constant (Fig. 2-7b), since the potential along the normal to the interface is constant. Thus, the component of the initial velocity directed along the normal to the interface is equal to the final velocity — i.e.,

$$v_1 \sin \alpha_1 = v_2 \sin \alpha_2. \tag{2-17}$$

Equation (2-17), which expresses the law for the refraction of an electron beam, is analogous to Eq. (2-16), which expresses the law for the refraction of a light ray. The difference consists in the fact that in Eq. (2-17) the velocities of the electrons v_1 and v_2 take the place of the refractive indices n_1 and n_2 that enter into Eq. (2-16). It follows from this that an electric field whose equipotential surfaces are similar in form to the surfaces of an ordinary optical lens will have the effect of a lens on an electron beam. It is true that it is impossible in practice to create an electron lens whose equipotential surfaces will be an accurate reproduction of the corresponding optical lens. However, electrostatic lenses are used in electron optics that are quite suitable for many practical applications.

We shall consider some of the most generally used electrostatic electron lenses. We should note first of all that the

homogeneous field already possesses innate focusing properties. Indeed, if there is no electric field between plane-parallel electrodes, electrons emanating from any point on the negatively charged electrode (cathode) move in radial directions and fall upon the whole surface of the positively charged electrode (anode) (Fig. 2-8a). But if a homogeneous electric field is created between these electrodes, the path of the electrons becomes curved. As a

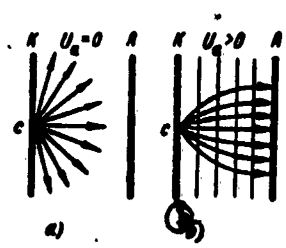

Fig. 2-8. Focusing action of a homogeneous electric field.

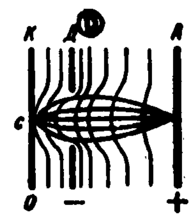

Fig. 2-9. Three-electrode focusing electrostatic lens.

result, a cone-shaped stream of electrons is formed in the space between the electrodes (Fig. 2-8b). Now the electrons do not fall on the whole surface of the anode, as their initial velocities cause them to do in the absence of a field, but on the surface of a comparatively small circle, whose radius is dependent on the magnitude of the field intensity and on the distance between the electrodes. The smaller this distance and the larger the field intensity, the smaller will be the radius of the circle. Thus, electrons can be collected into a stream of conical form under the influence of a homogeneous electric field — i.e., can be focused to a certain degree.

The three-electrode system is a better electrostatic lens; it consists of a cathode, diaphragm with a circular aperture, and an anode, all positioned parallel to each other (Fig. 2-9). If the anode is positive with respect to the cathode, and the dia-

phragm is negative with respect to the cathode, the equipotential surfaces will not be parallel to the surfaces of the cathode and the anode, but will take the form shown in Fig. 2-9. The electrons that emanate from point C of the cathode will move toward the anode under the influence of the accelerating field. But near the aperture of the diaphragm, the directions of increasing potential are toward the center of the aperture, and the lines of force of the electric field are directed from the center of the aperture to the edges of the diaphragm (Fig. 2-10). In such a case, the forces that act on the electron are directed toward the axis of the lens; and, in this direction, the field will accelerate the electrons. Under the influence of these forces, the path of the electrons will bend toward the axis of the lens and will land on the anode at one point. Thus, in the case considered, the diaphragm acts as a converging (focusing) lens.

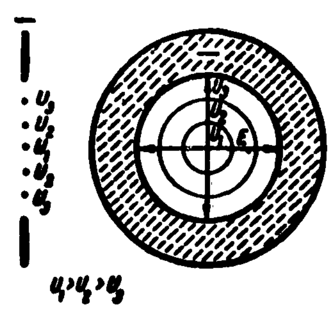

Fig. 2-10. Direction of the lines of force of the electric field of a negative diaphragm.

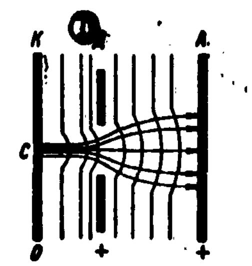

Fig. 2-11. Three-electrode electrostatic diverging lens.

If the potential of the diaphragm is positive and exceeds the potential of the surrounding space, the equipotential surfaces sag, as it were, toward the anode. Such a diaphragm acts as a diverging lens. This is explained by the fact that, in the region of the aperture, the direction of increasing potential is away

from the aperture, and the lines of force are directed toward the center of the aperture (Fig. 2-12). Consequently, the electrons are acted upon by a force which deflects them away from the axis of the lens. Therefore, the paths of the electrons diverge (Fig. 2-11).

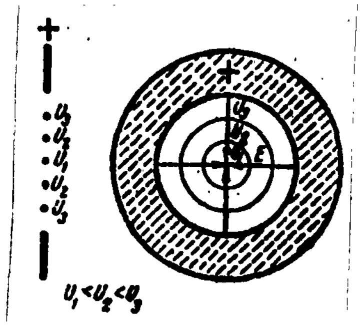

Fig. 2-12. Direction of the lines of force of the electric field of a positive diaphragm.

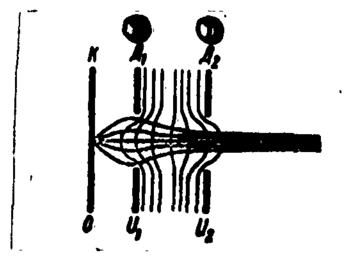

Fig. 2-13. Electrostatic lens with an accelerating field.

Electrostatic lenses with an accelerating field are very often utilized in the focusing of electron beams. These lenses consist of two diaphragms D_1 and D_2 and a cathode K (Fig. 2-13). Diaphragm D_1 usually has a small positive or negative potential. In order to create a strong accelerating field, the diaphragm D_2, as a rule, must have a high positive potential.

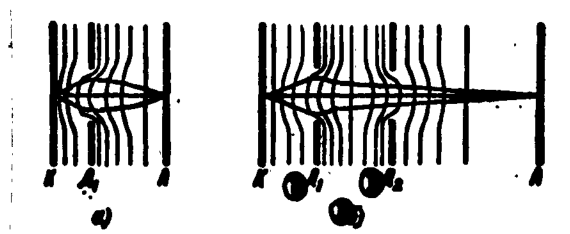

Fig. 2-14. Types of immersion lenses.

The equipotential surfaces of the electrostatic lens under discussion sag through the apertures of the first and second diaphragms. As a result, a focusing lens is formed in the region of the

aperture of the first diaphragm, and a diverging lens is formed in the region of the aperture of the second diaphragm. By varying the potentials of the diaphragms U_1 and U_2, the focal length of such a lens can be varied when the distance between the diaphragms and the cathode is constant. The electrostatic lens just discussed is used in cathode-ray tubes to obtain a small spot of light on the screen, the spot being the trace of a focused electron beam.

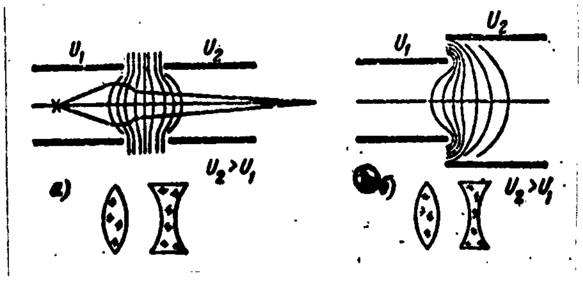

Fig. 2-15. Bipotential electrostatic lens and combinations of glass lenses analogous to them.

The following types of electrostatic lenses are those most frequently employed:

1. The immersion lens, which consists of one (Fig. 2-14a) or several diaphragms (Fig. 2-14b). The special feature of the immersion lens is the continuous varia[illegible]n of the potential of the electric field from the plane of the object to its image; this continuous variation is achieved by a suitable positioning of the diaphragms. An immersion lens can also be formed by two cylinders of the same or different diameters (Fig. 2-15a,b) with different potentials (bipotential lenses).

2. The unipotential lens, which is composed of three diaphragms (or three cylinders), of which the outer diaphragms are at the same potential, and the inner diaphragm is at a different potential (Fig. 2-16).

By utilizing electrostatic lenses, it is possible to focus an electron beam or to obtain a magnified image of the object by means of electronic beams just as a magnified image of an object or a focused light ray is obtained by means of optical (glass) lenses. This is explained by the fact that the formula for the focal length in light optics remains valid for electron optics:

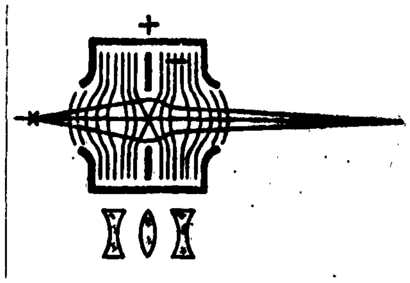

Fig. 2-16. Unipotential electrostatic lens and combination of glass lenses analogous to it.

$$\frac{1}{F}=\frac{1}{a}+\frac{1}{b},$$

where F is the focal length of the lens;

a is the distance from the object to the lens;

b is the distance from the lens to the image (Fig. 2-17).

The size of the image is found by the formula

$$B=A\frac{b}{a}$$

[when (b/a)>1, the image is magnified; when (b/a)<1, the image is diminished].

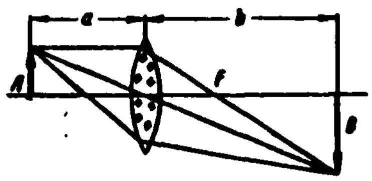

Fig. 2-17. Path of light rays when obtaining a magnified image.

2-3. THE MOTION OF AN ELECTRON IN A HOMOGENEOUS MAGNETIC FIELD.

In many electronic devices, the electron stream is controlled not only by means of electric fields, but also by magnetic fields. Magnetic fields exert a definite influence on the motion of electrons. This influence may be easily determined if use is made of the laws of the interaction of a magnetic field with an electric current.

Let us investigate the simplest case of the motion of an electron in a homogeneous magnetic field — i.e., in a field whose lines of force are straight, mutually parallel, and have the same density.

The force F with which a homogeneous field acts on an electron moving in it with a velocity v and creating an elementary current I' can be determined by Ampere's formula:

$$F = I'SH\sin\alpha, \qquad (2\text{-}18)$$

where S is the length of the segment (element) of the conductor on which a magnetic field with an intensity H acts; α is the angle between the direction of the magnetic field H and the direction of the elementary current I'.

If the electron, which has the charge $-e$, has covered the path S in time t, the elementary current created by the electron will be equal to:

$$I' = -\frac{e}{t}.$$

Substituting this value of the current in formula (2-18), we obtain:

$$F = -\frac{e}{t}SH\sin\alpha.$$

But S/t is simply the initial velocity of the electron v_0. Therefore, the force with which a homogeneous magnetic field acts on an electron moving in it with a velocity of v_0 is equal to:

$$F = -ev_0 H \sin\alpha. \qquad (2\text{-}19)$$

The direction of this force may be determined by the so-called "corkscrew rule": if the handle of a corkscrew is turned from H to v, as shown in Fig. 2-18, the corkscrew will move in (screw into) the direction of the force F. Thus, the force F is always perpendicular to the direction of the lines of force of the magnetic field and the direction of the velocity of the electron.

If the electron is moving rectilinearly with an initial velocity v_0 and enters a homogeneous magnetic field, it will be deflected from its original rectilinear path by the force F and will move along a complex path whose form is determined by the initial conditions of the motion of the electron.

We shall discuss a few cases of the motion in a homogeneous magnetic field given various initial conditions.

a) At the initial moment of time, the velocity of the electron v_0 is equal to zero. Substituting this value of the initial velocity into formula (2-19), we find that the force F = 0. Consequently, the magnetic field does not act on a motionless electron.

b) The initial velocity of the electron is not equal to zero and is directed along the lines of force of the magnetic field. In this case, the angle α between the direction of the lines of force and the direction of the initial velocity of the electron is equal to 0 or 180° and, consequently, $\sin(0^\circ) = \sin(180^\circ) = 0$; therefore the force F is also equal to zero. Thus, the magnetic field does not act on the electron in this case either. It follows

from this that the electron will move along the lines of force or against them, changing neither the magnitude nor the direction of its initial velocity (Fig. 2-19).

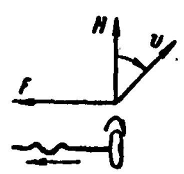

Fig. 2-18.
Corkscrew rule.

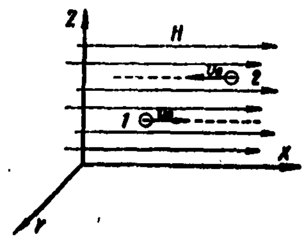

Fig. 2-19. The path of the motion of an electron in a homogeneous magnetic field if: 1 - $\alpha = 0^{\circ}$ and 2 - $\alpha = 180^{\circ}$.

c) The initial velocity of the electron is not equal to zero and is directed perpendicularly to the lines of force of the magnetic field — i.e., the angle is equal to 90 or 270°. Given these initial conditions, $\sin\alpha = \pm 1$. Substituting these values of $\sin\alpha$ into Eq. (2-19), we obtain:

$$F = \pm e v_0 H. \tag{2-20}$$

The force F is always perpendicular to the direction of the velocity of the electron. Therefore, the force does not do work in displacing the electron and, consequently, does not change its kinetic energy. The only action of the force F is to impart an acceleration of constant magnitude $F/m = \pm e v_0 H/m$ to the electron, which, like the force F, is directed perpendicularly to the velocity of the electron — i.e., is a centripetal acceleration. But the centripetal acceleration of the electron can also be expressed by v_0^2/r — i.e.,

$$\frac{e v_0 H}{m} = \frac{v_0^2}{r}.$$

It follows from this that the radius of curvature of the path of the electron is

$$r=\frac{mv_0}{eH}. \quad (2\text{-}21)$$

Since all the factors on the righthand side of Eq. (2-21) are constant, the radius of the path of the electron is also constant. It follows from this that an electron entering a homogeneous magnetic field with an initial velocity v_0, directed perpendicularly to the lines of force of the field, will perform a circular motion in the field under the influence of a force F; the circle will lie in a plane perpendicular to the lines of force of the field (Fig. 2-20). In this motion, the force F, the centripetal acceleration, and the initial velocity of the electron will change only in direction; they will not change their magnitude.

d) The initial velocity of the electron is directed at a certain angle α to the lines of force of the field. In this case, the velocity may be resolved into a component $v_{0\parallel}$, directed along the lines of force, and a component $v_{0\perp}$, directed perpendicularly to the lines of force of the field. Under the influence of the former component, the electron will move, as in case "b", uniformly and rectilinearly along the lines of force; under the influence of the latter component, the electron will describe, as in case "c", a circle lying in a plane perpendicular to the lines of force of the field. Under the influence of both components, the electron will describe a helix, whose radius may be determined from Eq. (2-21) by substituting the perpendicular component for v_0:

$$r=\frac{mv_{0\perp}}{eH}. \quad (2\text{-}22)$$

The axis of the helix will be directed along the lines of force of

the field (Fig. 2-21).

Thus, a magnetic field differs from an electric field in that the former acts on an electron only if the electron cuts the lines of force of the field during the course of its motion. In this case, only the path of the electron is altered; the kinetic energy and the velocity remain constant.

Let us consider the motion of an electron in a homogeneous magnetic field from the point of view of electron optics.

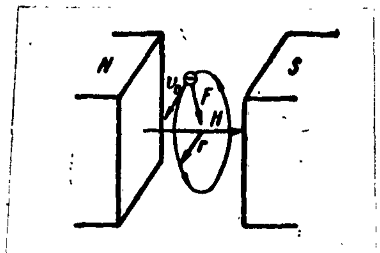

Fig. 2-20. The motion of an electron in a homogeneous magnetic field (initial velocity of electron directed perpendicularly to the lines of force of the field.

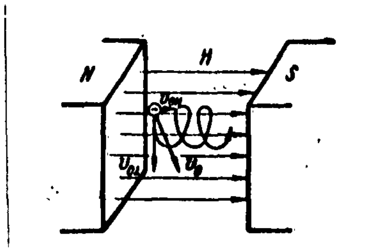

Fig. 2-21. The motion of an electron in a homogeneous magnetic field (initial velocity of electron v_0 directed at an angle α to the lines of force of the field).

Let a homogeneous magnetic field be created by means of the passage of an electric current I through a sufficiently long solenoid (Fig. 2-22). If, from point 1, the electrons travel at various angles to the axis of the solenoid, several of them — i.e., those whose velocities are in a direction coinciding with that of the axis of solenoid — will move rectilinearly along the axis of the solenoid. But the overwhelming majority of the electrons which are traveling at various angles α to the axis of the solenoid will move, as we know, in helices whose radii are determined by the corresponding components of velocity $v_{0\perp}$ perpendicular to the axis of the solenoid:

$$r = \frac{mv_{0\perp}}{eH}.$$

The time that it takes the electron to move through a full turn of the helix, the turn being of length $2\pi r$, may be found by dividing the length of the loop by the linear velocity of the electron:

$$\boxed{\tau = \frac{2\pi r}{v_{0\perp}} = \frac{2\pi m}{eH}.}$$

We see from the formula that this time is determined by the

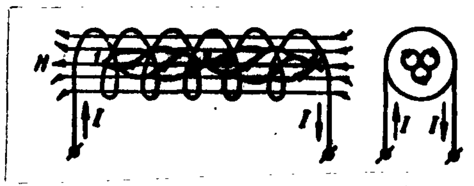

Fig. 2-22. Focusing action of a homogeneous magnetic field.

intensity of the magnetic field H and is dependent neither on the velocity of the electron $v_{0\perp}$, nor on the angle α. If the components of the velocity directed along the axis are approximately the same in magnitude, the electrons coming from point 1 will describe the turns of helices of various radii, but in the same time — i.e., all electrons coming from point 1 will collect at point 2 as at a focus after time τ; after this, they will again diverge and collect at point 3 after time τ, and so on. Points 2, 3, etc., are direct, unmagnified images of point 1 (multiple focusing). Thus, an electron stream can be focused by means of a solenoid. However, such a focusing method is inconvenient because in order to obtain a homogeneous magnetic field, sufficiently long solenoids (coils) must be employed.

2-4. THE MOTION OF AN ELECTRON IN NONHOMOGENEOUS MAGNETIC FIELDS. MAGNETIC ELECTRON LENSES.

Electron streams are much more often focused by means of so-

called "short" magnetic lenses. These lenses are coils of small length by means of which nonhomogeneous magnetic fields can be obtained that possess axial symmetry and fully satisfactory focusing properties.

We shall consider the focusing action that the nonhomogeneous magnetic field of a short lens has on an electron stream. At point A, which is near the coil and at a distance of r from the X axis, the intensity of the nonhomogeneous field can be resolved into two components: a component H_x, directed along the X axis, and a radial component H_r. Let electrons leave the cathode from point 1 with initial velocities of v_0, approximately equal in magnitude, and let the direction of the velocities make a small angle α with the X axis (paraxial electrons). Since the electrons cut the lines of force of the magnetic field during the course of their motion, the radial component of the field will act on the electrons with a force $F_\beta = ev_0H_r$. This force is perpendicular to the plane formed by the component H_r and the initial velocity v_0; and, consequently, it imparts a lateral acceleration to the electrons, as a result of which the electrons are deflected from their original direction to the side by an angle β. By virtue of this, the electrons acquire a velocity component v_β in the lateral direction. Owing to this component, the electrons will move in a circle of radius r around point B. At the same time, owing to this same velocity component v_β, a force $F_r = ev_\beta H_x$ directed toward the X axis will begin to act on the electrons; under the influence of this force, the paths of the electrons will bend toward the X axis and will again converge at the point 2, which lies on the X axis to the right of the coil. This point is a focus of the short magnetic lens. The image obtained of the object will be rotated by

an angle β with respect to the object. Thus, the short magnetic lens makes it possible to obtain a real, magnified image of the object, rotated by an angle β. The angle of rotation of the image is dependent on the velocity of the electrons and the intensity of the magnetic field. The magnification obtained by means of a short magnetic lens is determined, as in the case of the ordinary optical lens, by the ratio b/a.

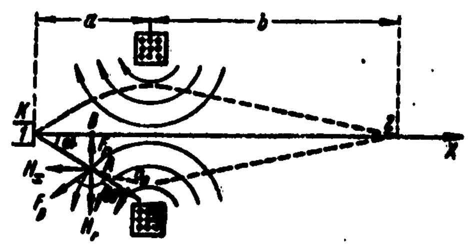

Fig. 2-23. Focusing action of the nonhomogeneous magnetic field of a short lens.

The focal length of magnetic electron lenses may be evenly varied by changing the magnitude of the current in the coils. This is an advantage not possessed by optical lenses.

Magnetic electron lenses are subdivided into three classes according to the magnitude of the focal length, which is dependent on the extent of the field. These classes are: long lenses, in which the magnetic field is homogeneous from the object to its image; medium lenses, in which the field is nonhomogeneous, and its extent is comparable to the focal length; and short lenses, in which the extent of the magnetic field is small in comparison with the focal length. In order to obtain strongly magnified, high-quality images, magnetic lenses must have a small focal length and produce minimal distortions.

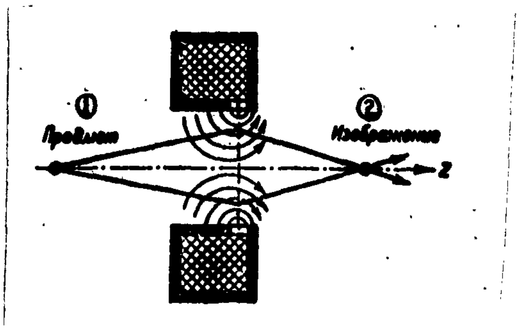

Fig. 2-24. Encased magnetic coil. 1) Object; 2) image.

In order to obtain a magnetic lens with a small focal length, it is necessary to create a magnetic field of large intensity at a small section along the axis of the lens; this intensity is dependent on the number of ampere-turns — i.e., on the product of the number of turns by the coil current. If the dimensions of the coil are constant, the intensity of the field can be increased by increasing the current. However, when the current is increased, the intensity grows not only at the center of the coil, but also at its edges. In increasing the current, the amount of the increase is limited by the heating up of the coil. The heating of the coil can be diminished if the cross-section of the conductor is increased, but increasing the cross-section means increasing the dimensions of the coil and the so-called "crawling away" of the field, which consists in the fact that the intensity of the field increases more at the edges of the coil than at its center, as a consequence of which the focal length increases.

At the present time, the focal length is decreased not only by increasing the field as much as possible, but also by "compression" of the field along the axis of the coil. The field is compressed by screening the coil — i.e., by placing it in an iron case in whose inner side there is a slit. Since the magnetic

field of iron is much larger than for air, the intensity of the magnetic field will be much higher near the slit than in a coil that is not screened (Fig. 2-24).

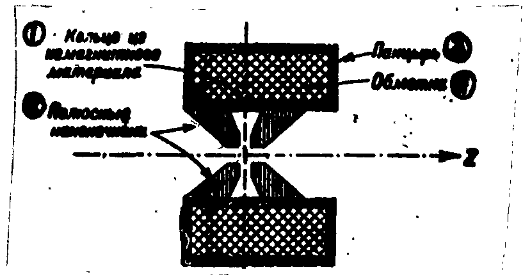

Fig. 2-25. Encased magnetic coil with pole pieces.
1) Ring of nonmagnetic material; 2) pole pieces; 3) case; 4) winding.

Still more decrease in the focal length can be achieved by the use of so-called "pole pieces," which are placed inside the coil opposite the slit (Fig. 2-25); this decreases the internal diameter of the coil and further compresses the field. Pole pieces are made out of ferromagnetic alloys and are composed of two cones which are opposite poles of a powerful electromagnet.

Chapter Three

ELECTRON EMISSION

3-1. TYPES OF ELECTRON EMISSION.

In Chapter One, it has been shown that electrons may be led out from a solid into the space surrounding it, if a certain additional energy is imparted to them by any means. Regardless of the way in which the additional energy is imparted to the electrons, or of the conditions that may be set up for the escape of electrons from the solid, the process of liberation of electrons from a solid into the space surrounding it is called electron emission. Depending upon the method used to impart additional energy to the electrons of the solid, or upon the conditions set up to provide liberation of electrons from the solid, various types of electron emission are distinguished.

At the present time, the following types of electron emission are known:

1. Thermionic emission in which additional energy is imparted to the electrons by heating the solid.

2. Field emission (electrostatic emission) occurring as a result of external electric fields ·ting up high field strength at the surface of the solid.

3. Secondary electron emission, occurring upon bombardment of the solid by a stream of electrons (called primary), or any other charged particles. When this happens, a portion of the kinetic energy of the primary electrons (or particles) is imparted to the electrons of the solid.

4. Photoelectric emission, occurring as a result of the effect of a stream of radiant energy, which is partially converted into

kinetic energy of the electrons of the solid.

3-2. THERMIONIC EMISSION. THERMIONIC-EMISSION EQUATION.

The phenomenon of thermionic emission may be observed by assembling a simple electrical circuit (Fig. 3-1), consisting of a source of direct current B_n *, a galvanometer G, a switch P, and a glass envelope B, within which a vacuum has been set up. A metal plate A (anode) and a metal filament K (cathode) are located in the glass envelope, a certain distance apart; the filament is heated to a high temperature by passing an electric current through it. If the plate A is connected in series with the galvanometer G to the positive side of the filament battery B_n, the galvanometer needle will be deflected, indicating the existence of a weak electric current in the plate circuit, and, consequently, in the evacuated space between the plate and the cathode as well. This current always appears with a heated filament and a positive potential on the plate. With the filament cold, the galvanometer indicates no current with either a positive or negative potential on the plate. From this it follows that the hot filament is a source of negative electrical charges, since they are only able to move under the action of a positive potential on the plate.

The true nature of the negative electrical charges produced by a hot filament in a vacuum has been known for some time. In 1889, Thomson established the fact that the ratio of the charge of particles given off by a hot filament to their mass is precisely equal to the charge-mass ratio of the electron.

Thus, it has been established that a hot filament yields (emits) electrons; as the electrons move in the vacuum as a result

*[B_H - B_n - B_{nakal} - $B_{filament}$]

of the action of an accelerating electric field toward the plate, which is at a positive potential with respect to the filament, they create an electric current opposite in direction to the motion of the electrons.

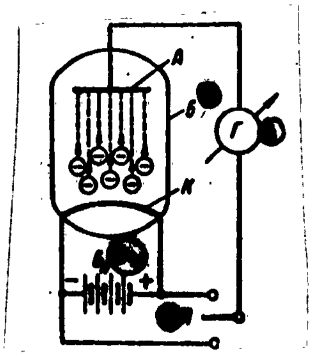

Fig. 3-1. Basic circuit for observing thermionic emission.
A) Metal plate (anode); K) hot metal filament (cathode);
G) galvanometer;
B) glass envelope;
B_n) filament battery;
P) switch.

At a high filament temperature, certain electrons leave the filament with quite high initial velocities. Thus, a thermionic-emission current may be observed in the plate circuit with zero potential, or even with small negative potentials, on the plate, since the fast electrons may overcome weak retarding fields, and reach the plate.

3-3. THERMIONIC EMISSION AS A FUNCTION OF TEMPERATURE AND WORK FUNCTION.

At room temperature, metals emit almost no electrons. The explanation for this lies in the circumstance that the kinetic energy of the overwhelming majority of free electrons in the metal is less than the total work function. With an increase in temperature, the velocity and kinetic energy of a considerable portion of electrons increase. Those electrons whose energy becomes equal to the total work function may leave the metal and go out into the surrounding space. The higher the temperature of the heated metal, the greater the number of electrons escaping its surface, and the greater the thermionic-emission current. Thus, the magnitude of the thermionic-emission current bears a fixed relationship to the temperature of the heated metal.

For pure metals, the thermionic-emission-current density I'_e

depends upon the temperature T in accordance with the equation obtained theoretically by Richardson and Dushman:

$$I'_e = AT^2 e^{-\frac{b_0}{T}}. \qquad (3\text{-}1)$$

This equation indicates that the current density I'_e varies with changes in metal temperature T according to an exponential law. The constants A and b_0 entering into the equation have a specific physical interpretation.

The constant b_0 is the temperature equivalent of the work function, equal to 11,600°K. It is defined by the equation $b_0 = e\varphi_0/k$, where e is the charge of an electron; $k = 1.38 \cdot 10^{-16}$ erg · degree^{-1} – the Boltzmann constant; $e\varphi_0$ is the effective work function. Thus, the constant b_0 depends upon the work function for the metal.

TABLE 3-1.

1) Металл	Температура плавления, 2) °K	3) A, $a/см^2 \cdot град^2$	$b_0 = \frac{e\varphi_0}{k}$, °K	φ_0, эв 4)
5) Вольфрам	3 655	60—100	52 400	4,6
6) Молибден	2 895	55	48 700	4,16
7) Тантал	3 300	60	47 500	4,10
8) Торий	2 130	70	39 200	3,35
9) Барий	977	60	24 500	2,52
10) Цезий	300	162	21 000	1,81

1) Metal; 2) melting point, degree K;
3) A, amp/cm^2 · degree^2; 4) φ_0, ev;
5) tungsten; 6) molybdenum; 7) tantalum;
8) thorium; 9) barium; 10) cesium.

The quantity A is a constant whose theoretically computed value is:

$$A = 60.2 \text{ amp} \cdot \text{cm}^{-2} \cdot \text{degree}^{-2}.$$

Experience has shown, however, that for real metals, A has differing values, larger and smaller than 60.2 (Table 3-1).

A does not even remain constant for exactly the same metal, but changes as its surface condition varies.

The discrepancy between theory and experiment has been eliminated by modern electronic theory, which in its derivations of the law of thermionic emission takes into account the shape of the potential barrier, and the velocity of the electrons approaching it.

Equation (3-1) indicates that as the temperature rises and the work function drops, the thermionic-emission-current density increases.

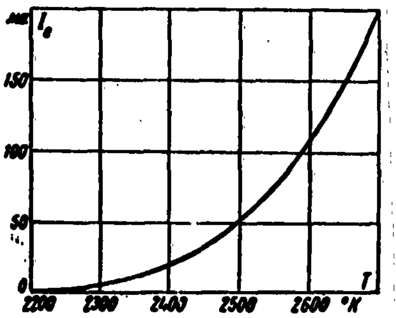

Fig. 3-2. Thermionic-emission current of a tungsten cathode as a function of temperature.

Figure 3-2 is a graphical representation of the law of thermionic emission for a tungsten cathode having a surface of 0.145 cm^2. This exponential curve indicates that as the temperature of the tungsten rises from 2200 to 2400°K, the thermionic-emission current rises slightly, while when the temperature increases from 2400 to 2700°K, the current rises very rapidly. The shape of the graph is accounted for chiefly by the exponential term $e^{(-b_0/T)}$ in the equation. It is easily calculated that for tungsten heated to 2500°K, a temperature increase of 1%, i.e., to 2525°, causes an increase in the T^2 term of about 2%, and in the $e^{(-b_0/T)}$ term of 20%.

The law of thermionic emission may be expressed graphically for any metal, if the constants A and b_0 are known for it. For this purpose, appropriate temperature values are taken, expressed in degrees absolute, and the current density I'_e in amp/cm^2 is computed according to Eq. (3-1). From the computed data, a graph of $I_e = f(T)$ can be plotted, with temperature values T in degrees absolute along

the X axis, and the total current I_s, in milliamperes, over the entire surface of the heated metal along the Y axis.

3-4. THERMIONIC EMISSION FROM ACTIVATED SURFACES.

Although several of the metals have low work functions, they still cannot be used in pure form to obtain any considerable emission owing to unsuitable physical properties, and chiefly owing to low melting point. In practice, it has proved possible to raise the working temperature of metals with low work functions somewhat by combining them with other, higher-melting metals, and as a result to increase the thermionic emission. For example, if a thin layer of thorium is applied to a tungsten surface, the temperature of the combination may be raised to a value close to the melting point of thorium. The working temperature of similar combinations of metals may be above the working temperature of the low-melting metal, and the work function lower than that of each of the metals taken in isolation.

Using the method described, it has proved possible to utilize the low work function of certain metals at higher temperatures, at which these metals would melt in pure form. For example, the melting points of tungsten and thorium equal 3655 and 2130°K respectively. With a layer of thorium on tungsten, the work function amounts to 2.63 ev, while at the same time, the work function of pure tungsten is 4.6 ev, and that of pure thorium 3.4 ev. Such thoriated tungsten gives large emission at a working temperature of about 1900°K over an extended period of time, which is impossible with pure thorium.

Barium and cesium are also sometimes used in addition to thorium for coating higher-melting metals. Such combinations are called "barium on tungsten," "thorium on tungsten," "barium on

nickel," etc. Thorium on tungsten is the most frequently utilized combination. Barium on tungsten and cesium on tungsten are less often used, since barium and cesium evaporate comparatively easily from the surface of the heated tungsten.

In addition to coatings of thin layers of low-melting metals on the surfaces of higher-melting metals, coatings are also used consisting of thin layers of oxides of metals having low work functions. In 1904 it was found that at comparatively low temperatures (on the order of $1000^{\circ}K$) very great thermionic emission was produced by oxides of the alkali-earth metals: barium, strontium, and calcium; these materials were applied to the surface of another metal and subjected to appropriate heat treatment. It was found later that as a result of heat treatment, free atoms of barium formed in the oxide layer; they then diffused throughout the thickness of the oxide layer and on its surface, and decreased the work function, thus causing a strong increase in thermionic emission.

TABLE 3-2.

1) Комбинация металлов	2) A, $а/см^2 \cdot град^2$	$b_0 = \frac{e\varphi_0}{k}$, °K	3) φ_0, эв
4) Торий на вольфраме (83%-ное покрытие)	2,8	34 150	2,94
5) Торий на вольфраме (100-ное покрытие)	3,0	30 500	2,63
6) Торий на тантале	0,5	29 200	2,52
7) Торий на карбиде вольфрама	$1,2 \cdot 10^{-2}$	17 400	1,49
8) Окись бария на вольфраме . .	$(0,2-20) \cdot 10^{-2}$	11 200	0,97
9) Смесь окиси бария с окисью строиция на никеле	$2,6 \cdot 10^{-3}$	9 350	0,81

1) Combination of metals; 2) A, amp/cm^2·degree^2; 3) φ_0, ev; 4) thorium on tungsten (83% coating); 5) thorium on tungsten (100% coating); 6) thorium on tantalum; 7) thorium on tungsten carbide; 8) barium oxide on tungsten; 9) barium oxide-strontium oxide mixture on nickel.

Table 3-2 gives the values of the work functions and other thermionic-emission constants for several of the most common metals and oxides in combination with other metals. The data of the table indicate that the values of the constants depend not only on the nature of the coating utilized, but also upon the degree to which the base metal is coated.

The process through which large and stable thermionic emission is obtained by applying a thin layer of some other metal having a low work function to the surface of a base metal is called activation of the surface of the base metal. Metals having a high work function and relatively high melting point (tungsten, molybdenum, nickel, etc.) are activated. Alkali and alkali-earth metals are used as the activating metals; they should be electropositive with respect to the base metal.

The magnitude of the thermionic-emission current from pure and activated surfaces may be computed in accordance with Eq. 3-1. It should be kept in mind, however, that in the case of activated surfaces, the value of the constants A and b_0 depend very heavily upon the degree and uniformity with which the active layer covers the surface of the base metal.

3-5. EFFECT OF ACCELERATING ELECTRIC FIELD UPON THERMIONIC EMISSION. FIELD EMISSION.

In the absence of an external electric field, the only electrons that are able to leave the heated metal are those that have a kinetic energy equal to or greater than the work function at the instant they cross the metal-vacuum interface. In this case, the thermionic emission is defined by the well-known equation

$$I_t = AT^2 e^{-\frac{\varphi_0}{kT}}, \qquad (3-2)$$

where I'_t is the thermionic-emission current density in the absence of an external electric field.

An external accelerating field always causes a certain increase in emission, even through the temperature of the metal remains constant. In the given case, the accelerating field tends to counteract the forces holding the electrons in the metal, and thus decreases the work function and, consequently, increases emission. Thus, in the presence of an external accelerating field the thermionic-emission equation (3-2) becomes inaccurate.

The practical utilization of thermionic emission is always connected with external electric fields. Consequently, for a more accurate determination of the work function and the magnitude of thermionic emission, it is necessary to take into account the effect of the external electric field.

Let us consider in basic terms the mechanism by which electrons escape from the surface of a hot metal in the presence of an external electric field.

Let an electron escape from a metal whose surface is assumed to be flat. In direct proximity to the surface, there are attractive forces acting on the electron due to the ions located near the point at which the electron has left. To the degree that the electron moves away from the surface, it begins to feel the effect of the positive charge induced by it, and concentrated in the metal at the point of mirror reflection of the receding electron. In this case, the distance x increases, while the force acting on the electron decreases. For graphical purposes, however, it is more convenient to deal not with the mirror-reflection-force variation, but the electrical potential φ corresponding to it. In Fig. 3-3, the change in the potential of the mirror-reflection force is repre-

sented by curve 1, indicating a sharp drop in potential with distance from the metal surface; this means that a retarding electrical field (potential threshold) acts upon electrons receding from the metal.

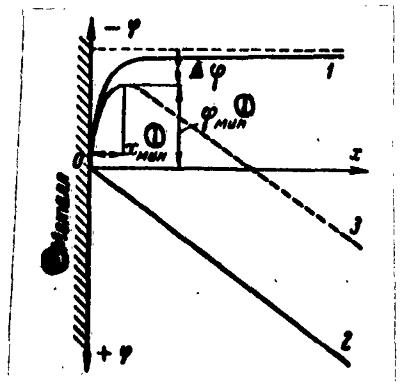

Fig. 3-3. Conversion of a potential threshold into a potential barrier under the influence of an external accelerating field.
1) Min; 2) metal.

Curve 2 represents the distribution of the potential of an external electric field where the electrodes have plane-parallel surfaces; it follows from this curve that the potential of the external field rises in proportion to distance. For electrons departing from the metal, this electric field will be accelerating. In the last analysis, an electron departing from the metal is subject at every point of space to the action of a resultant potential, which takes the form of a sum of the potentials due to the mirror-reflection force and the external electric field.

The resultant-potential distribution is shown in Fig. 3-3 by a dashed line (curve 3). Taking at each point of curve 3 the tangent of the slope, representing the field strength E, and multiplying it by the electron charge e, we may compute the force with which the resultant electric field acts upon an electron departing from the heated metal, and construct a graph of its variation. This force retards the motion of the electron along the path segment from $x = 0$ to $x = x_{min}$, and accelerates it along the segment $x > x_{min}$. Thus, where there is an external accelerating field, those electrons escape the heated metal whose kinetic energy is sufficient to overcome the retarding force. These electrons proceed a distance of up

to $x = x_{min}$ in a retarding field (potential barrier), and then enter the accelerating field, while in the absence of an external field, the same electrons would travel a greater distance in the retarding field than $x = x_{min}$, and the potential difference is greater by an amount $\Delta\varphi$ than in the presence of an external field. In the last analysis, in the presence of an external accelerating field, an electron does an amount of work in leaving the metal that is less by an amount $e\Delta\varphi$. In this case, the work function will equal not φ_0, as in the absence of the external field, but $\varphi_0 - \Delta\varphi$. Thus, the equation for thermionic emission in the presence of an external electric field takes the form:

$$I'_{tE} = AT^2 e^{-\frac{e(\varphi_0 - \Delta\varphi)}{kT}} = AT^2 e^{-\frac{e\varphi_0}{kT}} e^{\frac{e\Delta\varphi}{kT}}, \qquad (3\text{-}3)$$

where I'_{tE} is the thermionic-emission-current density, depending upon the temperature and upon the external electric field.

The theoretical computation gives the following expression for the magnitude of the emission with an external field strength E:

$$I'_{tE} = I'_t e^{4.4\frac{\sqrt{E}}{T}}. \qquad (3\text{-}4)$$

If there is no external field (E = 0), Eq. 3-4 becomes Eq. 3-2:

$$I'_{tE} = I'_t = AT^2 e^{-\frac{e\varphi_0}{kT}}.$$

which defines the thermionic emission from a heated metal in the absence of an external field.

It follows from the considerations given that if the field strength at the emitting surface is sufficient for complete compensation of the retarding effect of the potential threshold, then even at low temperatures, it might be possible to obtain large

electron emission. Calculation shows that a field strength on the order of 10^8 v/cm is required to compensate the potential threshold. Even where the field strength is on the order of 10^6 v/cm, i.e., 100 times less, large electron emission is observed from cold surfaces.

The mechanism by which electrons escape from the surface of a cold metal is more complicated than that involved in thermionic emission, and cannot be explained by the fact that under strong accelerating fields, a metal loses only those electrons whose kinetic energy is equal to or greater than the energy w_a. On the contrary, with field emission, electrons leave the metal surface with all kinds of velocities, low and high. In order to explain this phenomenon, we must abandon the simplified picture of an electron as a sphere whose motion in an electric field obeys the laws of classical mechanics, and take account of wave properties. From the point of view of wave mechanics, it is permissible for even low-energy electrons to pass through the potential barrier (tunnel effect). In this case, the narrower the potential barrier, the greater the probability for electrons to pass through it.

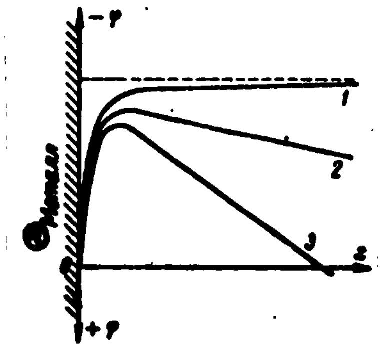

Fig. 3-4. Change in shape of potential barrier as a function of magnitude of external field.
1) $E = 0$; 2) $E > 0$;
3) $E \gg 0$; 4) metal.

In the absence of an external field, the potential barrier may be represented by curve 1 (Fig. 3-4), in the presence of a weak external field, by curve 2, and with a strong external field, by curve 3. Clearly, as follows from curve 1, in the absence of an external field, the only electrons that can penetrate the potential barrier are those whose energies by chance prove equal to

or greater than the total energy w_a, the probability of which is low, since the metal is cold. Thus, the potential barrier proves to be opaque for the vast majority of electrons.

In the presence of a weak accelerating field, the width of the potential barrier decreases, but its height, as indicated by Fig. 3-4, remains nearly unchanged. In this case, the potential barrier will not be completely opaque, but all the same, a very small number of electrons will penetrate the barrier. Where there are strong accelerating fields, the barrier becomes considerably narrower (curve 3), and its transparency rises sharply. Thus, the greater the field strength near the surface of a cold metal, the narrower the potential barrier becomes, and consequently, the greater the field emission. Of course, for the same field strength, cold emission will be greater for metals with low work functions.

In order to determine the field-emission current density accurately, it is necessary to know, first, the fraction of the surface of the metal which actually gives off electrons, and, second, the actual value of the external field strength, which may vary sharply owing to unavoidable nonuniformities in the surface of a real metal (points, protuberances, etc.).

With thermionic emission, as we know, a metal gives up only the fastest electrons. The departure of these electrons from the metal is accompanied by a drop in its temperature. Conversely, with field emission, the metal loses chiefly slow electrons, and a temperature drop is not observed.

3-6. PHOTOELECTRIC EMISSION. PHOTOEMISSION LAWS.

The phenomenon of photoelectric emission or photoemission was discovered experimentally in 1887. Detailed quantitative studies of photoelectric emission were carried out in 1889 by A. G. Stoletov,

who illuminated various metal plates in his experiments (zinc, copper, nickel, etc.) with bright light from an arc lamp (Fig. 3-5). If the plate investigated was at a positive potential, no current appeared when it was illuminated. Every time that the plate was at a negative potential, when it was illuminated, a sensitive galvanometer connected into the circuit showed a weak current, which appeared and vanished together with the light. This current is called a photoelectric current, or simply a photocurrent, and the electrons forming it, liberated from the metal under the action of the light — photoelectrons. Thus, in photoelectric emission, light energy is converted into electrical energy.

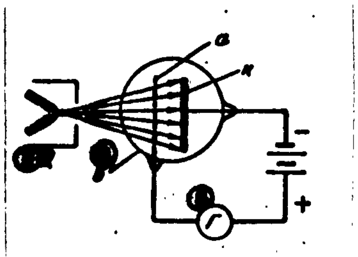

Fig. 3-5. Diagram of Stoletov's experiments. a) Electrode in form of anode grid; k) metal plate investigated (cathode); B) evacuated glass envelope; D) arc lamp; G) galvanometer.

The following basic laws have been established experimentally for photoelectric emission:

1. Stoletov's law. When a metal surface is illuminated by a light having precisely the same spectral composition, the number of electrons escaping in unit time from the surface of the metal (the photoelectric-emission current), is proportional to the intensity of the incident light, i.e., to the light flux:

$$i_a = kF, \qquad (3\text{-}5)$$

where i_a is the photoelectric-emission current; F is the light flux; k is a coefficient characterizing the sensitivity of the given surface to light, which is defined as the ratio of the photoelectric-emission current i_a to the light flux F, and is expressed in μa/lumen: $k = i_a/F$, μa/lumen.

There are two types of sensitivity: integral and spectral. By integral sensitivy we mean the sensitivity of the surface under investigation to white light (not broken down into the spectrum), and by spectral sensitivity, we mean sensitivity to monochromatic light, i.e., to light of one and only one wave length. Spectral sensitivity is arbitrarily designated by k_λ.

The linear relationship specified by Eq. 3-5 between the number of photoelectrons escaping from the metal and the magnitude of the light flux is explained by the fact that an increase in the intensity of the light is nothing else but an increase in the number of quanta of radiant energy making up this flux, and making possible the escape of electrons from the surface of the metal.

2. Einstein's law. The maximum energy of the electrons leaving the surface of a metal in photoelectric emission is a linear function of the frequency of the light incident upon the surface of the metal, and is independent of its intensity.

The correctness of this law is easily seen. Let us designate the energy of a valence electron of a metal by W. If this electron receives a quantum of radiant energy $h\nu$ in size, its energy will then equal $W = h\nu$. Clearly, those electrons may leave the metal whose total energy $W + h\nu$ is equal to or greater than the energy W_a required for an electron to escape from the metal, i.e., $W + h\nu \geqslant W_a$. After leaving the metal, some of these electrons will have kinetic energy left over, and its magnitude may be determined from the equation

$$\frac{mv^2}{2} = W + h\nu - W_a.$$

The greatest kinetic energy $(mv^2/2)_{max}$ will be possessed by those electrons leaving the surface of the metal which while still in the metal, before absorbing the energy of the quantum, had a maximum

energy W_1, which they retain upon moving to the surface. On this basis, we may write:

$$\left(\frac{mv^2}{2}\right)_{\text{макс}} = W_i + h\nu - W_a = h\nu - e\varphi_0, \qquad (3\text{-}6)$$

where $W_a - W_1 = e\varphi_0$, the effective work function of an electron leaving metal (in ergs), and $h\nu$ is the magnitude of a quantum of radiant energy from momochromatic light.

Equation 3-6 is called the Einstein equation, expressing the second law of photoelectric emission, and indicating that the maximum energy with which electrons leave a given metal depends solely upon the frequency of the incident light, and is independent of its intensity, since an increase in light intensity is equivalent to an increase in the number of light quanta which only increase the number of electrons escaping the metal, without affecting their energy.

Equation 3-6 makes it possible to determine the so-called threshold frequency at which and below which photoelectric emission is not observed at all. In fact, if a metal is illuminated with monochromatic light, and the wave length is gradually increased (or the frequency decreased), at a certain value $\lambda_0 = c/\nu_0$, the energy of the quantum will prove equal to the work function: $h\nu_0 = e\varphi_0$. In this case, Eq. 3-6 takes the form:

$$\left(\frac{mv^2}{2}\right)_{\text{макс}} = 0.$$

Consequently, at a wave length $\lambda_0 = c/\nu_0$, photoelectric emission ceases. This wave length is called the "threshold" for photoelectric emission, since with a further increase in wave length, the energy of the quanta proves to be less than the work function, and the electrons absorbing the quanta cannot escape from the metal. At wave lengths greater than λ_0, therefore, photoelectric emission is not observed

even with very high light intensity.

Naturally, as temperature rises, electrons will appear in the metal with energies greater than the maximum energy W_1, and their escape from the metal requires only the absorption of quanta with low energies corresponding to the long wave length. As temperature increases, therefore, the photoelectric-emission threshold shifts toward longer wave lengths. In actuality, at room temperature the number of such electrons is small, and almost no shift in the photoelectric-emission threshold is observed.

Utilizing the relationships $\lambda_0 = c/\nu_0$ and $\nu_0 = e\varphi_0/h$, it is possible to establish a connection between the work function of a metal φ_0 and the threshold wave length λ_0:

$$\varphi_0 = \frac{ch}{e\lambda_0} = \frac{12\,400}{\lambda_0},$$

where φ_0 is measured in electron volts, and λ_0 in angstroms.

The expression obtained indicates that the lower the work function of a metal, the lower should be the energy of the quanta that is necessary to produce photoelectric emission, i.e., photoelectric emission appears at longer light wave lengths. Consequently, with metals having low work functions, photoelectric emission should be observed at long wave lengths. For example, the photoelectric-emission threshold of alkali metals (potassium, cesium, etc.), lies in the long wave portion of the spectrum.

A graphic picture of the photoelectric-emission threshold and the wave length regions in which photoelectric emission takes place is given by the so-called spectral response for photoelectric emission, which takes the form of a curve representing spectral sensitivity k_λ as a function of the light-flux wave length λ: $k_\lambda = f(\lambda)$. According to the formula, photoelectric-emission spectral response is classified

into two types: normal and selective. If, from λ_0 onward, the sensitivity of the metal surface increases as wave length decreases (Fig. 3-6, curve a), the photoelectric emission is called normal. As a rule, normal photoelectric emission is observed in pure metals used in the form of comparatively thin layers. In practice, together with normal photoelectric emission, we observe the so-called selective photoelectric emission, in which the curve of the sensitivity-wave length function, i.e., the spectral response, has a single, or several, maxima, indicating that the given surface has increased sensitivity to light in a narrow wave length band (λ_{izb},* Fig. 3-6, curve b). Selective photoelectric emission appears especially strongly in comparatively thin, specially treated films of alkali metals.

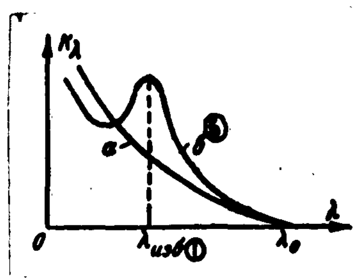

Fig. 3-6. Photoemission spectral response.
a) Normal photoemission response; b) selective photoemission response;
1) λ_{izb}.

It should be kept in mind that of the total number of light quanta incident upon the surface of a metal, only a small number of them will be absorbed by electrons that may later leave the metal. This is explained by the fact that the majority of metals have a large reflection coefficient (about 85 per cent). Thus, a large number of quanta are reflected from the surface of the metal. In addition, of the quanta falling upon the metal, a considerable number penetrate

* [изб - izb - izbiratel'nyy - selective]

a long way into the metal (200-300A) and electrons can escape from a metal from depths not exceeding several tens of angstroms. As a result, a considerable amount of the light quanta penetrating the metal are absorbed by the metal and their energy goes to heat the metal.

Finally, those electrons which absorb light quanta do not always leave the metal, owing to unfavorable motion conditions (for example, velocity directed to the interior of the metal). In the final analysis, for every thousand quanta incident upon the metal surface, approximately one electron escapes from the metal, i.e., of the thousand quanta incident, one quantum proves to be "effective." The ratio of the number of effective quanta to the total number incident upon the surface is called the quantum efficiency or quantum output for photoelectric emission, and for pure metals amounts to a fraction of a per cent.

3-7. SECONDARY ELECTRON EMISSION.

A very important quantity characterizing secondary electron emission is the secondary-emission coefficient σ, representing the ratio of the number of secondary electrons n_2 liberated from a metal to the number of primary electrons n_1 bombarding the metal, or the ratio of the secondary-emission current I_2 to the primary-electron current I_1:

$$\sigma = \frac{n_2}{n_1} = \frac{I_2}{I_1}. \quad (3-7)$$

If σ and I_1 are known, then it is easy to use Eq. 3-7 to determine the secondary-emission current I_2:

$$I_2 = \sigma I_1. \quad (3-8)$$

The value of the secondary-emission coefficient σ depends both on the properties of the metals and the condition of its surface, and on the

velocity of the primary electrons.

It has been established experimentally that at high enough primary-electron velocities, the number of secondary electrons may considerably exceed the number of primary electrons.

Fig. 3-7 gives curves showing the variation of σ with a change in the velocity of the primary electrons for several of the metals most commonly used in electronic devices. The curves show that as the primary-electron velocity increases, the secondary-emission coefficient rises, reaching a maximum (σ_m), and then dropping with a further increase in velocity; the maximum values for the secondary-emission coefficient for different metals corresponds to different values of the primary-electron velocities. Thus, for example, for tungsten the maximum of σ corresponds to a primary electron velocity of 800v, while for molybdenum, it is 400v.

The appearance of maxima on the curves is connected with the fact that primary electrons moving at different velocities penetrate into the metal to different depths. At low velocities, primary electrons cannot penetrate deep into a metal; in this case, the majority of secondary electrons are now formed by primary electrons reflected from the metal. As velocity increases, the primary electrons penetrate deeper into the metal — the higher the velocity, the greater the penetration. Within the metal, the primary electrons yield their energy to the electrons of the metal, which, having obtained this additional energy, are in a state permitting them to escape from the metal. The greater the velocity of the primary electrons, the greater the number of electrons of the metal receiving additional energy, and the more electrons escaping from the metal. As a result, there is a maximum on the curve, corresponding to a specific critical velocity of the primary electrons, and, consequent-

ly, a specific critical depth h_{kr}* for their penetration into the

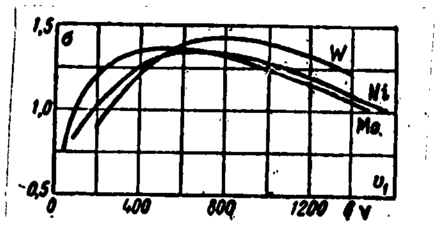

Fig. 3-7. Secondary-emission coefficient σ as a function of primary-electron velocity v_1 for tungsten, molybdenum, and nickel.

metal (Fig. 3-8a). With a further increase in velocity, however, the primary electrons penetrate so deep into the metal ($s > h_{kr}$), that the electrons of the metal, obtaining additional energy at such a depth, in moving toward the surface collide with other electrons and atoms, thereby losing a portion of their energy; the remaining energy is no longer adequate for escape from the metal. As a result, the secondary-emission coefficient begins to drop with a further increase in the depth of penetration of the primary electrons.

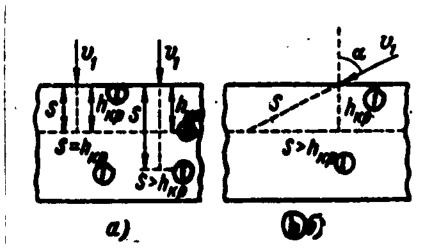

Fig. 3-8. As the angle of incidence α and velocity v_1 of the primary electrons increase, their path s in the metal lengthens. 1) kr.

Secondary-electron emission depends not only on the magnitude, but also upon the direction of the primary-electron velocity. It has been established experimentally that if primary electrons strike the

*[kp - kr - kriticheskiy - critical]

surface of a metal along the perpendicular, the secondary emission will be maximum. To the degree that there is an increase in the angle α between the normal to the surface and the direction of the primary-electron velocity, the secondary emission increases. On theoretical grounds, it is assumed that in this case the primary electrons, striking the metal, traverse a distance $s > h_{kr}$ (Fig. 3-8b), but to a depth with respect to the surface not exceeding h_{kr}. In traversing the longer path, they have a larger number of collisions with electrons of the metal. As a result, a large number of electrons of the metal obtain additional energy from the primary electrons at a depth not exceeding h_{kr}. In moving to the surface, these electrons of the metal lose almost none of their energy, since the probability of a collision with other particles along this short path to the surface is low.

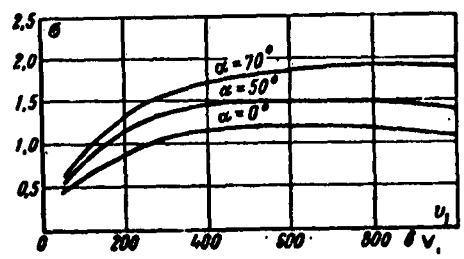

Fig. 3-9. Coefficient of secondary emission as a function of primary-electron velocity v_1 for various values of the angle of incidence α.

At such angles, therefore, a large number of electrons of the metal, obtaining additional energy from the primary electrons, pass through the surface of the metal. As a result, the secondary-electron emission, and, consequently, the secondary-emission coefficient as well, reach a maximum at large angles α.

Figure 3-9 gives curves for σ as a function of v_1 and of α. For pure metals, the maximum secondary-emission coefficient is

smaller. For example, for nickel $\sigma_m = 1.35$, for tungsten $\sigma_m = 1.41$. For certain dielectrics the secondary-emission coefficient is greater than for metals. Thus, for mica, $\sigma_m = 3$, for glass $\sigma_m = 3.1$ (Table 3-3). This is explained by the fact that in dielectrics, the conductance of the electrons is very small, and thus the secondary electrons, in moving to the surface, lose nearly none of their energy through interaction with conduction electrons.

If a thin layer of other metals, especially alkali metals, is applied to the surface of a pure metal, with a layer of oxides of these metals, then the coefficient σ for these complex surfaces is about ten times higher in value than for the pure metals.

TABLE 3-3.

Металлы, диэлектрики (1)	σ_m
(2) Цезий	0,72
(3) Никель	1,35
(4) Молибден	1,37
(5) Вольфрам	1,41
(6) Тантал	1,53
(7) Слюда	3
(8) Стекло	3,1

1) Metals, dielectrics;
2) cesium; 3) nickel;
4) molybdenum; 5) tungsten;
6) tantalum; 7) mica;
8) glass.

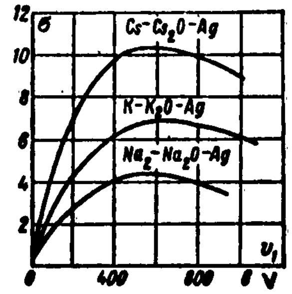

Fig. 3-10. Secondary-emission coefficient as a function of primary-electron velocity for composite surfaces.

Figure 3-10 gives curves showing σ as a function of primary-electron velocity v_1 for three different composite surfaces: σ increases, evidently, owing to a decrease in the work function. It is not always the case, however, that surfaces with low work functions have high secondary-emission coefficients. For example, the work function of tungsten is high ($\varphi_0 = 4.6$ v), but the secondary-emission coefficient is also high in comparison with other metals ($\sigma = 1.5$). If a tungsten surface is exposed to oxygen, φ_0 increases to 9.25v,

while the coefficient σ remains nearly unchanged.

Especially great secondary-electron emission is obtained where there are 10^{-4} to 10^{-5} cm thick dielectric films on the metal, covered in turn by a thin film of alkali metal (the Malter effect). Under these conditions, secondary electrons escape not only owing to primary electrons, but also owing to strong electric fields occurring in a dielectric layer. Primary electrons, striking the film of alkali metal, cause secondary emission. But the departure of the secondary electrons from the film is accompanied by the formation of positive charges on the dielectric surfaces; owing to the poor conductance of the dielectric, the positive charges remain unneutralized for a long period of time. As a result, a strong electric field appears between the positive charges and the metal, in the thin layer of dielectric; owing to the small thickness of the layer, the field strength reaches 10^6 v/cm. Such strong fields continue to draw electrons from the metal through the dielectric layer even after primary-electron bombardment has ceased and thus this phenomenon has great persistence.

Secondary-electron emission from thin dielectric films covered with a thin film of alkali metal, in view of the great persistence, has received almost no application, despite the enormous value of the secondary-emission coefficient, which sometimes reaches values of 1000.

The phenomenon of secondary-electron emission has been successfully employed in several types of electronic devices for current amplification. In other electronic devices, however, spontaneously appearing secondary emission results in strong distortion of the characteristics of the device, and thus plays a negative role.

Chapter Four

HOT CATHODES FOR ELECTRONIC DEVICES

4-1. CLASSIFICATION OF HOT CATHODES.

In the majority of electronic devices, a stream of free electrons is provided by thermionic emission, i.e., from a heated metal electrode, which is normally connected to the negative side of the circuit, and is called the cathode. The most convenient method of heating a cathode is to pass an electric current through it.

Cathodes emitting free electrons when heated are called thermionic or hot cathodes.

Modern hot cathodes are subdivided into three varieties on the basis of the form of the emitting surface:

1. Cathodes made from pure metals or alloys (simple cathodes).
2. Film cathodes, i.e., metal cathodes having a one-atom thick film of another electropositive metal on the surface.
3. Oxide cathodes.

In addition to the division based upon physical properties, hot cathodes are classified into two large groups according to the method by which they are heated. The first group comprises the so-called directly heated (filamentary) cathodes, and t second the indirectly heated (heater-type) cathodes first suggested in 1918 by academician A.A. Chernyshev.

Filament-type cathodes are heated to the working temperature by passing an electric current directly through the cathode itself, which emits the free electrons. Indirectly heated cathodes are heated by a special electrical heater, which, as a rule, is insulated from the cathode itself either by an evacuated space, or by a special insulating layer which has high heat resistance and excellent mechanical properties.

In certain electronic devices, the emission current required for their operation is obtained from so-called cold cathodes, which, unlike hot cathodes, do not require a special heater. A description of cold cathodes will be given when we take up electronic devices in which they are utilized.

In recent years, reports have appeared of a new type of highly efficient cathode, which is very similar in its properties to cold cathodes, and as a result has been called a new type of cold cathode. It utilizes the phenomenon of the persistence of secondary-electron emission from a thin layer of dielectric (MgO) on nickel. The new cathode is brought into operation with an expenditure of power (the so-called starting power), which is less by a factor of 10 than the power required for heating a standard cathode. Still less power is sufficient to maintain emission during operation of a cathode. In the first samples of tubes using this cathode, the emission current amounted to several dozens of microamperes. In order to provide normal tube operation, however, plate voltages of not less than 300v are required.

4-2. CHARACTERISTICS AND PARAMETERS OF HOT CATHODES.

In choosing the most suitable mode of operation of a hot cathode, it is possible to utilize the dependence of the emission current on the cathode temperature. It is, however, quite difficult to carry out a direct measurement of the cathode temperature in a manufactured tube, and especially under operating conditions. In practice, it proves more convenient to measure and to monitor not the temperature, but the voltage (or current) used to heat the cathode, employing for this purpose standard metering equipment. Assigning various values of heating voltage or current, it is possible to measure the values of emission current corresponding to them, and on the basis of this data to construct a graph showing emission current as a function of cathode heat-

ing voltage (or current).

The curve showing thermionic-emission current as a function of cathode heating voltage $I_e = f(U_n)$ or of heating current $I_e = f(I_n)$ is called the emission characteristic of a hot cathode.

Another extremely important characteristic of a cathode is the heating characteristic, expressing the cathode heating current I_n as a function of the heating voltage E_n.

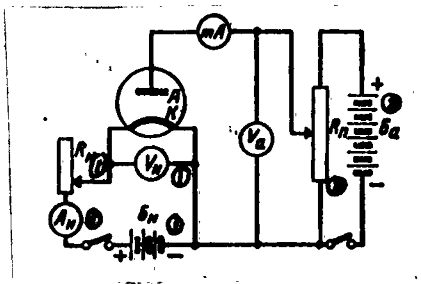

Fig. 4-1. Circuit for finding emission and heating characteristics of a cathode.
1) n; 2) B_a; 3) p.

The emission and heating characteristics of a cathode may be obtained experimentally by using a circuit consisting of a plate loop and a filament loop (Fig. 4-1). The filament circuit consists of the series-connected rheostat R_n, the filament battery B_n, the ammeter A_n, and the cathode K. The rheostat R_n serves to vary the heating current and voltage, which determine the cathode temperature. The voltage drop across the cathode (cathode heating voltage) is measured with the aid of a voltmeter V_n, and the heating current with the aid of ammeter A_n.

The plate loop consists of a voltage source B_a, potentiometer R_p*, and voltmeter V_a, with the aid of which the various values of potential difference between plate and cathode are established (plate

*[n - p - pitaniya - supply.]

voltage U_a), and the milliammeter mA, used to measure the emission current. The filament loop and plate loop have a common point at which the potential is arbitrarily assumed to equal zero. The potentials of the remaining points in the circuit are calculated with respect to this common point.

When a voltage U_a has been set for the plate, various values of heating voltage (or current) U_n are applied, and the values of plate current I_a corresponding to these heating values are read off from the milliammeter connected into the plate circuit.

To find an emission curve, it is necessary to make a proper choice of the plate voltage U_a. If the plate voltage is large enough, all of the electrons escaping from the cathode will reach the plate. In this case, the plate current I_a, measured with the aid of the milliammeter, will equal the emission current I_e. If, however, the plate voltage is not large enough, a certain portion of the electrons coming off from the cathode, affected by the repulsive forces of the electrons that have previously left, may return to the cathode, and the milliammeter will show a plate current I_a that is less than the emission current I_e. With a proper choice of plate voltage (so that $I_a = I_e$), on the basis of the measured values of plate current, we construct the curve $I_e = I_a = f(U_n)$ or $I_e = f(I_n)$, i.e., the emission curve.

With the aid of this same circuit, it is possible to find the heating characteristic of the cathode. To do this, the voltmeter V_n is used to establish various values of cathode heating voltage U_n, and the corresponding heating-current values I_n are read off from ammeter A_n. From the data obtained, a heating curve is constructed.

The heating and emission curves may be found simultaneously, if the emission current I_e is small in comparison with the heating current

I_n. If the emission current is on the order of the cathode heating current, the ammeter connected into the cathode heating circuit will measure a current larger or smaller than the cathode heating current I_n, depending upon whether it is connected into the positive or negative side of the heating circuit. In this case, the plate-voltage supply circuit must be opened and the heating curve derived separately.

To choose the operating mode of a cathode, it is convenient to plot the emission and heating curve on the same graph, and to use them to determine the voltage (or current) for heating at which the cathode yields the emission required for operation of the tube.

Figure 4-2 shows sample emission and heating curves. The cathode emission curve essentially represents a graph of the law of thermionic emission (Fig. 3-2), with the sole difference that instead of plotting the temperature T along the X axis, we plot the heating voltage or current, which determines the cathode temperature. The curve of Fig. 4-2 shows that the emission curve begins not from zero, but from a certain value of heating voltage U_n, at which the number of electrons leaving the weakly heated cathode is small, but all the same constitutes a noticeable emission current. As the heating voltage increases, the temperature of the cathode rises, and the emission current begins to increase slowly, but at an ever higher and higher rate.

It also follows from the graph of Fig. 4-2 that the heating curve is not linear. This is explained by the rise in the resistance of the cathode when it becomes hot: as the resistance of the cathode rises, the heating current increases more slowly than the heating voltage.

In addition to choosing the operating mode for the cathode, the emission and heating curves make it possible to determine the cathode parameters, i.e., the quantities most characteristic of the cathode,

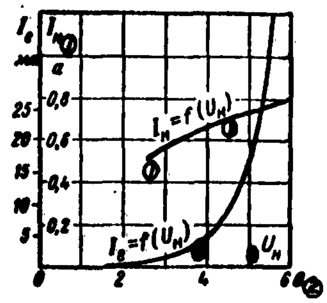

Fig. 4-2. Emission [$I_e = f(U_n)$] and heating [$I_n = f(U_n)$] curves for a cathode.
1) n; 2) v.

which determine its properties.

The basic parameters of a hot cathode are the specific emission (emission-current density) I'_e, the specific heating power P'_n, the effectiveness (efficiency) H, and the life time (service life) of the cathode τ.

Specific emission. The specific emission of the cathode is defined as the magnitude of emission current obtained from one cm^2 of cathode surface at the working temperature. Knowing the total emission current I_e over the entire working surface of the cathode S_n, it is possible to determine the specific emission I'_e:

$$I'_e = \frac{I_e}{S_n} \text{ [amp/cm}^2\text{]}$$

Specific emission is also determined from the thermionic-emission equation:

$$I'_e = AT^2 e^{-\frac{b_0}{T}}.$$

It follows from this that the specific emission is a function of temperature, the properties of the material, and the state of the cathode surface.

In practice, the specific emission may be determined if we know the geometry of the cathode (total working surface), and the total emission current, which may be found from the emission curve.

Specific heating power. In order to heat a cathode to the working temperature, it is necessary to apply a specific amount of power to it. During the first instant that heating current is turned on, this power is used for:

a) raising the cathode temperature; b) the thermal radiant energy supplied by the cathode to the surrounding space; c) loss of heat to the cathode supports and insulators; d) creation of the initial kinetic energy of the electrons leaving the cathode.

When an equilibrium condition sets in, in which the applied power becomes equal to the power lost by the cathode, the cathode temperature ceases to rise, and power ceases to be expended upon heating of the cathode. From this moment onward, the power applied to the cathode is expended solely in thermal radiation, in loss of heat to supports and insulation, and in the energy removed from the cathode by the electrons given off by it. Here, the main portion of the heating power goes into the radiation of heat into the space nearby, and this depends upon the cathode temperature and its surface properties. The magnitude of this power is determined by the Stefan-Boltzmann law:

$$P'_n = \varepsilon\sigma T^4, \qquad (4\text{-}1)$$

where P'_n is the power radiated by $1 cm^2$ of cathode surface at a temperature of T^oK; $\sigma = 5.72 \cdot 10^{-12} watt/cm^2 \cdot degree^4$ – the Stefan-Boltzmann constant; ε is a radiation coefficient depending upon the surface properties of the cathode. For an ideal black body, $\varepsilon = 1$, for all real surfaces, ε is less than 1.

The majority of electrons leave the surface of a hot cathode with

nonzero initial velocities, and, consequently, with nonzero initial kinetic energies. Each of these electrons removes an energy from the cathode equal to its initial kinetic energy. This energy removed by the electrons, however, is slight and amounts to about 5 per cent of the total heating power.

The loss of heat to supports and insulation depends upon the cathode structure and the method by which it is mounted. For a directly heated cathode, mounted on comparatively thin supports, the energy lost through the heat carried off is slight. For indirectly heated cathodes, having considerable contact surfaces with more massive supports and insulators, the loss may amount to 20 per cent of the power drawn by the cathodes.

It is normally assumed that all of the power applied to a cathode is used up in radiation and, consequently, is found from Eq. (4-1). In this case, the energy losses due to cooling by supports and insulation and by giving off of electrons are neglected.

The heating power expended per $1cm^2$ of cathode working surface is called the specific heating power for the cathode:

$$P'_н = \frac{P_н}{S_н} \; [вт/см^2]. \quad *$$

It is desirable to expend as little heating power as possible in maintaining the working temperature of a cathode; the only way to do this is to decrease cooling by supports and insulators.

<u>Cathode efficiency</u>. The cathode efficiency N is the ratio of the total thermionic-emission current I_e to the total power P_n expended in heating the cathode at the working temperature:

$$H = \frac{I_e}{P_н} \; [ма/вт]. \qquad (4\text{-}2)$$

---

*[вт - vt - vatt - watt.]

In other words, the cathode efficiency is the magnitude of the emission current per watt of heating power. Evidently, the higher the emission current obtained per watt used in heating the cathode, the greater the efficiency, and the more economical the cathode. Since for a given material, the specific emission and specific heating power depends solely upon the cathode temperature, the efficiency also is a function of temperature alone. Actually, if into the expression for efficiency (4-2) we substitute the values of the total emission term I_e, expressed in terms of the specific emission I'_e, and of the total heating power P_n, expressed in terms of the specific power E'_n, we obtain an equation determining the nature of the dependence of efficiency upon temperature:

$$H=\frac{I_e}{P_n}=\frac{I'_e S_n}{P'_n S_n}=\frac{AT^2 e^{-\frac{b_0}{T}}}{\varepsilon\sigma T^4}=\frac{A}{\varepsilon\sigma}T^{-2}e^{-\frac{b_0}{T}}.$$

As the temperature rises, the term $e^{-b_0/T}$ rises more rapidly than the term T^{-2} drops. It follows from this that in order to increase cathode efficiency, it is necessary to increase its working temperature, i.e., it is more advantageous to utilize a cathode at a higher temperature.

The last of the very important cathode parameters is the lifetime, defined by the total continuous cathode operating time during which it provides the necessary emission. The lifetime of simple directly heated cathodes is normally defined as the operating time during which emission current does not drop below 80 per cent of its initial value. For various types of electron tubes, cathode life fluctuates from several tens to several thousands of hours, depending upon the type of tube and its function.

Evaporation of cathode material at high operating temperatures plays a decisive role in the life of simple cathodes. Here, more in-

tense evaporation occurs at the more highly heated portions. As a result, at the sections, the cathode diameter decreases more than at other ones. The decrease in diameter, in the last analysis, leads to destruction or overheating of the cathode. The total operating time of the cathode until it has completely burned out is called the total cathode service life. Owing to a drop in emission current, however, caused by a decrease in the working surface owing to a decrease in cathode diameter, long before the cathode has completely burned out, it actually proves to be unsuitable for service. Thus, the so-called useful life has been established for simple cathodes; it is the total operating time of the cathode during which the emission current at constant heating voltage drops to a previously arbitrarily established value. Frequently, the useful life is determined by the total operating time of the cathode during which its diameter in the very hottest portion decreases by 10 per cent of the initial value.

The life of film and oxide cathodes is basically determined by evaporation of the active layer, resulting in a loss of emission by the cathode, so that it becomes unsuitable for service. In such cathodes, the active layer of electropositive materials evaporates especially vigorously even upon slight overheating. The life of activated cathodes is normally determined experimentally, since theoretical determinations are difficult.

4-3. STRUCTURE OF HOT CATHODES.

Hot cathodes of modern electronic devices are classified according to the method by which they are heated, and on the basis of their structure, into two types: directly heated cathodes and indirectly heated cathodes. Directly heated cathodes may in turn be simple or activated (film, oxide). Structurally speaking, simple cathodes do not

differ from activated cathodes. Indirectly heated cathodes are manufactured solely in the oxide form.

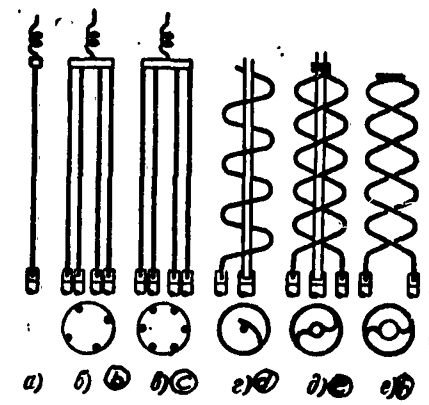

Fig. 4-3. Construction of directly heated cathodes for cylindrical electrode systems.

The simplest arrangement is used with directly heated cathodes, which take the form of a metal filament (conductor), shaped in various manners depending upon the structure and dimensions of the other electrodes utilized in the given device. The simplest design is the straight filament (Fig. 4-3a), mounted on two metal supports which are also utilized to lead the heating current to the cathode. If the filament is long, it is made in one shape or another depending upon the structure of the other elements (Fig. 4-3b, c, d, e, f). Cathodes of this type are utilized in devices having a cylindrical structure, and are arranged either coaxially with the other electrodes, or with axes parallel at equal distances from the axes and from each other. If the filament is very long, it takes the form of a spiral (Fig. 4-3d), or a double spiral (Fig. 4-3e, f) with a support (Fig. 4-3d, e), or without a support (Fig. 4-3f).

In flat designs, the cathode is formed from several parallel filaments arranged in a single plane (Fig. 4-4a, b), or in the form of a single loop or several loops taking the form of inverted letters

V or W (V-shaped and W-shaped), also arranged in a single plane (Fig. 4-4c, d, e). Such cathodes are mounted on two or several current-conducting supports and maintained in tension by a single or several springs. It is necessary to apply tension to a filament in order to eliminate sagging resulting from lengthening upon heating.

Simple directly heated cathodes are preferably made of tungsten, and only in isolated instances of tantalum or niobium.

Cathode supports may be of various designs, depending upon the dimensions and shape of the cathode, and also upon its operating conditions in the tube. The basic requirement applying to supports is that they not only provide for the application of heating current to the cathode, but that they also provide the required mechanical strength of the cathode mount.

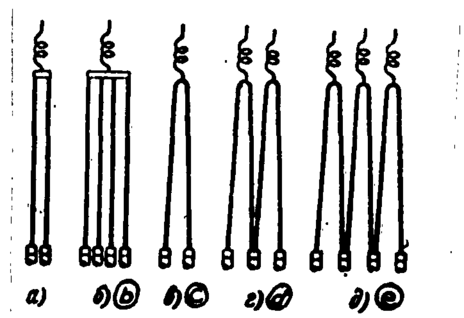

Fig. 4-4. Filaments for planar electrode system.

The major drawback to directly heated cathodes is the fact that when they are supplied with alternating current, the cathode temperature changes at double the frequency of the heating current, causing substantial changes in emission current, owing to which apparatus using tubes with such cathodes are subject to the appearance of low-frequency noise (at twice the AC frequency). Thick cathodes heat up and cool down more slowly, and thus their temperature and, consequently, emission current as well change less with heating-current

variations than do thin cathodes. Relatively thick cathodes, however, may be utilized only in power tubes. Only thin cathodes, naturally, may be used in low-power tubes.

More complicated in construction are heater-type cathodes which consist of the cathode itself, emitting the electrons, and a heater, with the aid of which the cathode is heated to the working temperature. The most common form of heater-type cathode is a hollow metal cylinder (Fig. 4-5a). In several electronic devices, flat rectangular or oval cathodes are used (Fig. 4-5b, c), or elliptical cathodes (Fig. 4-5d). In order to obtain sufficiently narrow electron beams, flat can-type cathodes of various heights and diameters are used (Fig. 4-5e, f), with emitting surfaces in the form of flat cylinders.

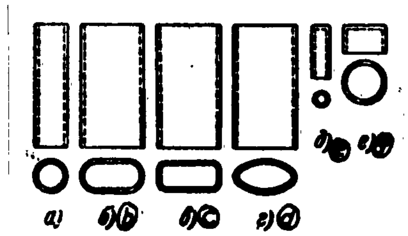

Fig. 4-5. Heated cathode designs.

The construction of heaters for indirectly heated cathodes is determined by the construction of the cathodes themselves, in which the heaters are to be used. The basic types of heater designs are shown in Fig. 4-6. Loop-type or folded heaters with a small number of loops (Fig. 4-6 a, b), are utilized for low heater voltages (up to 4v). Folded heaters with a large number of loops (Fig. 4-6c), or helical heaters (Fig. 4-6d, e) are utilized for high heater voltages (up to 30v).

In order to eliminate the effect of the magnetic field set up by the heater current, so that the electron-emission current will not

be affected, and in order to provide for uniform heating of the cathode, double-helical heaters are utilized (Fig. 4-6f).

At high heater voltages, where long thin conductor is used, the heater is made in the form of a double-helix from conductor that has been previously been made helical (Fig. 4-6g). In cases where the helices are made of thin wire, ceramic rods are used to give them greater strength (they are also called supports), as are tungsten dowels, previously insulated with alundum, or alundum tubes with the heater spirals inserted within or slipped over these tubes (Fig. 4-6h), depending upon the structure of the cathodes themselves.

For end-fire cathodes, the heaters are made in the form of flat (Fig. 4-6i) slightly stretched (Fig. 4-6j) or short cylindrical (Fig. 4-6k) helices.

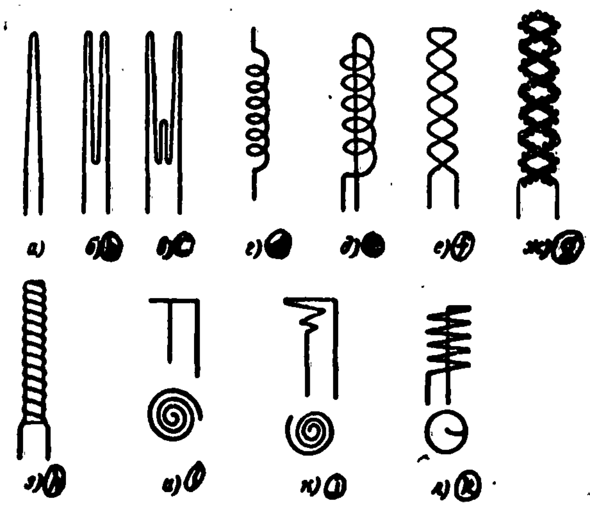

Fig. 4-6. Designs of various types of heaters for indirectly heated cathodes.

Nickel is normally used as the cathode material; an oxide layer is applied to the nickel surface. Since the working temperature of an oxide cathode is comparatively low, nickel proves to be a completely satisfactory material with respect to its properties for manufacturing heaters. In order to maintain the working temperature of the cathode,

however, the heater should be at a higher temperature. Thus, higher-melting metals — tungsten and its alloys — are used in manufacturing heaters. The heater plays no direct part in creating the thermionic emission of the cathode.

An advantage of heater-type cathodes is that they are equipotential devices, since the value of the potential at any point on the surface of the cathode is practically identical. Owing to large heat losses, however, the efficiency of heater-type cathodes is lower than the efficiency of directly heated cathodes. The thermal inertia of a heater-type cathode is large, and when the cathode is in service this plays a positive role, guaranteeing a constant temperature, and plays a negative role when the cathode is turned on, since it requires 15 to 25 sec to heat the cathode up. The high thermal inertia of heater-type cathodes makes it possible to utilize alternating current for heating, which is very advantageous and convenient in tube operation.

4-4. PURE-METAL CATHODES. DESIGNING A TUNGSTEN CATHODE.

As far as physical properties are concerned, the most suitable metals for manufacturing simple cathodes are tungsten, tantalum, and niobium, of which the most widely utilized is tungsten. Tantalum and niobium are utilized infrequently.

The wide-spread utilization of tungsten for manufacturing cathodes is explained by its high melting point ($3655^{\circ}K$), satisfactory ductility and malleability, necessary for manufacturing thin conductors, and excellent constant emission in time even under considerable ion bombardment, under which the operation of cathodes of other types becomes unstable.

The operating temperature of a tungsten cathode ranges from 2400-$2700^{\circ}K$. In this temperature range, the emission-current density varies substantially from 0.1 to 1.5 amp/cm^2, and the efficiency from

2 to 16 ma/watt.

Thicker cathodes, 1-1.5mm in diameter, as a rule, are heated to higher temperatures. Thin cathodes, with diameters of several hundredths of a millimeter, are heated to temperatures of 2400-2500°K, since at higher temperatures, even weak atomization of the tungsten causes a considerable decrease in the diameter of a thin filament, and this in turn leads to a further increase in temperature, and still more rapid vaporization of the tungsten. Thus, thin tungsten cathodes are very sensitive to even slight overheating, which results in a noticeable decrease in the emitting surface and, consequently, in the life as well.

The life of a simple cathode is defined in general form by the equation

$$\tau = \frac{\delta}{M}(r_0 - r_\tau)[\text{sec}]$$

where M is the amount of material evaporating in one sec from each cm^2 of cathode surface with the working temperature unchanged (in grams); δ is the density of the cathode material; r_0 is the cathode radius when service begins $(t = 0)$; r_τ is the cathode radius at the end of service $(t = \tau)$.

If we assume that the life equals the time over which the cathode radius decreases by 10 per cent, i.e., $r_\tau = 0.9r_0$ and, consequently, $r_0 - r_\tau = 0.1r_0$, we obtain the formula first found by V.I. Volynkin:

$$\tau = \frac{0,1\delta}{M} r_0 [\text{sec}] \qquad = \frac{0,1\delta}{3\,600M} r_0 [\text{hr}] \qquad (4\text{-}3)$$

The value of M for various temperatures of the cathode is given in special tables. For tungsten, $\delta = 19$ and, consequently,

$$\tau = \frac{1,9}{3\,600M} r_0 = 2,64 \cdot 10^{-4} \frac{d_0}{M}. \qquad (4\text{-}3a)$$

It follows from formula (4-3) that the larger the cathode radius

r_0, the longer the life τ. For example, for a tungsten cathode of radius $r_0 = 0.004$ cm, at a working temperature of $T = 2400°K$ ($M = 0.44 \cdot 10^{-9}$ g/cm$^2 \cdot$ sec), the lifetime is:

$$\tau = \frac{1{,}9 \cdot 0{,}004}{3\,600 \cdot 0{,}44 \cdot 10^{-9}} = \frac{0{,}0076}{0{,}44 \cdot 36} \cdot 10^7 = 4\,790 \text{ [hr]}.$$

If the cathode radius is increased to $r_0 = 0.005$ cm, the life will equal:

$$\tau = \frac{1{,}9 \cdot 0{,}005}{3\,600 \cdot 0{,}44 \cdot 10^{-9}} = \frac{0{,}0095}{0{,}44 \cdot 36} = 10^7 \approx 6\,000 \text{ [hr]}.$$

If, however, the cathode radius r_0 remains unchanged, but the working temperature is raised, the lifetime will be greatly shortened. For example, with $r_0 = 0.004$ cm and $T = 2400°K$, the life will be 4790 hr. For a cathode of the same radius, but operating at a temperature of $T = 2500°K$ ($M = 2.03 \cdot 10^{-9}$), the life will be:

$$\tau = \frac{1{,}9 \cdot 0{,}004}{3\,600 \cdot 2{,}03 \cdot 10^{-9}} = \frac{0{,}0076}{36 \cdot 2{,}03} \cdot 10^7 = 1\,040 \text{ [hr]}.$$

i.e., the lifetime has been divided by a factor of 4.6.

The examples given indicate that in order to increase the life of a cathode, it is best to operate it at as low a temperature as possible, and as large a cathode wire as possible should be chosen. Moreover, the established operating mode for the cathode should be maintained as accurately as possible, since even inconsiderable changes in temperature cause sharp variations in cathode life.

Designing a tungsten cathode. As is clear from what has been said, all of the basic parameters for a pure-metal cathode depend solely upon its temperature. For an ideal cathode, i.e., one whose temperature is constant over the entire length, it is possible to establish a connection between the emission parameters and the cathode geometry.

The resistance of a cathode heated to a given temperature T is determined from the well-known expression

$$R_T = \rho_t \frac{4 l_\kappa}{\pi d_\kappa^2} = R' \frac{l_\kappa}{d_\kappa^2}, \qquad (4\text{-}4)$$

where ρ_1 is the resistivity of the cathode material, ohm·cm, at the given temperature; l_k is the length of the cathode; d_k is its diameter, cm.

Let $4\rho_t/\pi$ = R'. It is evident that R' is the resistance of a 1-cm long cathode that is also 1 cm in diameter. Such a cathode is called a unit cathode, and for such a cathode, it is possible, by experiment or by computation, to find the resistance, the power radiated by the side surface, and the emission current as functions of temperature for any pure metal.

If at a given temperature a heating current I_n flows through the cathode, a power

$$P_n = I_n^2 R_T = I_n^2 R' \frac{l_\kappa}{d_\kappa^2}. \qquad (4\text{-}5)$$

is involved.

It may be assumed that the entire heating power under steady-state conditions is used for radiation from the side surface of the cathode. The radiated power is

$$P_{\text{из}} = a\sigma T^4 \pi d_\kappa l_\kappa = P' d_\kappa l_\kappa, * \qquad (4\text{-}6)$$

where P' = $\pi a \sigma T^4$ is the power radiated from the side surface of a unit cathode, since the magnitude of its side surface is π [cm^2].

It is clear that the quantity P', for a given metal, depends solely upon temperature.

Equating the right sides of Eqs. (4-5) and (4-6) we obtain:

* [из - iz - izluchayemiy - radiated.]

$$I_{\text{н}}^2 R' \frac{l_{\text{к}}}{d_{\text{к}}^2} = P' l_{\text{к}} d_{\text{к}}, \tag{4-7}$$

whence, letting $\sqrt{P'/R'} = I'$, we have:

$$I_{\text{н}} = \sqrt{\frac{P'}{R'}}\, d_{\text{к}}^{3/2} = I' d_{\text{к}}^{3/2}. \tag{4-8}$$

The quantity I' is the heating current required to maintain a given temperature of a unit cathode. Just as in the case of the quantities P' and R', I' depends solely upon temperature. Equation (4-8) defines the connection between the cathode heating current and its diameter at a given temperature.

The connection between the cathode heating voltage and the cathode dimensions may be established on the basis of Ohm's law:

$$U_{\text{н}} = I_{\text{н}} R_T = I' d_{\text{к}}^{3/2} R' \frac{l_{\text{к}}}{d_{\text{к}}^2}. \tag{4-9}$$

Let U' designate the product I'R'. The quantity U' is the voltage across the ends of the unit cathode, setting up the current I' through it. Then

$$U_{\text{н}} = U' \frac{l_{\text{к}}}{\sqrt{d_{\text{к}}}}. \tag{4-10}$$

This expression is correct for a cathode whose temperature is constant over the entire length (ideal cathode). Practically speaking, for real cathodes, the sections located near the supports have temperature below the temperature of the central portion of the cathode, owing to cooling by the massive supports. Figure 4-7 shows the distribution of temperature along a cathode that is fastened at two points. The higher the maximum temperature of the cathode, the more uniform the temperature distribution along the length.

The working temperature of a cathode is considered to be its temperature at the very hottest spot. To provide such a temperature,

a heating current whose magnitude is determined according to Eq. (4-8) should flow through the cathode.

The voltage that must be applied to the ends of a real cathode in order to force the required heating current through it should be less than for an ideal cathode, since the cathode resistance is decreased due to the cooled sections. The magnitude of the heating voltage for a real cathode may be found by decreasing the value

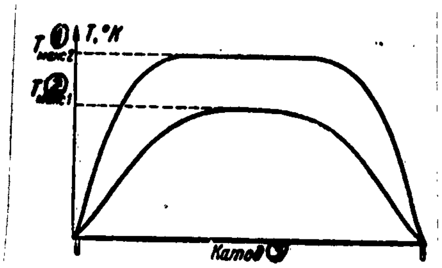

Fig. 4-7. Sample graph of the distribution of temperature over a directly heated cathode, fastened at two points.
1) $T_{max\ 1}$; 2) $T_{max\ 2}$;
3) cathode.

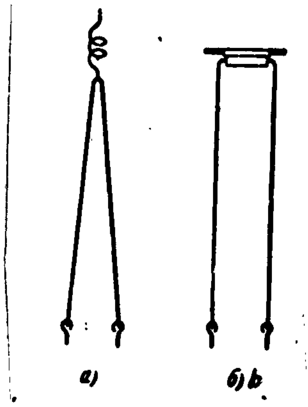

Fig. 4-8. Types of mounting for a loop filament.

calculated for an ideal cathode by the quantity $n\Delta U$, where n is the number of cooled sections of the cathode, and ΔU is a correction for the decrease in cathode resistance owing to the cooled portions. For tungsten cathodes, the quantity ΔU may be found on the basis of the following formula:

$$\Delta U = 0{,}00013(T_к - 400). \qquad (4\text{-}11)$$

Then the true voltage for heating a real cathode will be

$$U_к = U' \frac{l_к}{\sqrt{d_к}} - n\Delta U. \qquad (4\text{-}12)$$

For V-shaped filaments, n will equal three where the top of the loop is fastened to a thin spring (Fig. 4-8a). Where the top portion of the loop is fastened to heavy supports (Fig. 4-8b), n = 4.

In order to compute the emission from a real cathode, the concept of its effective length is introduced. The effective length of a real cathode is defined as the length of that ideal cathode whose temperature equals the working temperature of the real cathode, and which gives the same emission. Clearly, the effective length of a cathode is less than its actual length. The effective length of a cathode may be determined from the formula

$$l_{эф} = l_{к} - n\Delta l. \quad (4\text{-}13)$$

The correction Δl depends upon the temperature, diameter, and material of the cathode. For a tungsten cathode, the quantity l is computed in accordance with the formula

$$\Delta l = 2{,}72 \frac{\sqrt{d_{к}}}{F(T)}, \quad (4\text{-}14)$$

where F(T) is a temperature function. Determining the effective length of a cathode, we may easily compute the value of the emission by multiplying the value of the specific emission at the given temperature by the effective surface of the cathode, equal to $\pi d_k l_{eff}$:

$$I_e = I'_e \pi d_{к} l_{эф}.* \quad (4\text{-}15)$$

The parameters I', R', P', U' for a unit cathode, as well as other quantities needed for designing a cathode, the efficiency H, the specific emission I'_e, the evaporation rate M, and the function F(T), are given for tungsten in Table 4-1.

---

* [эф - ef - effektivnyy - effective - eff.]

TABLE 4-1.

T, °K	I' 1 $a/см^{3/2}$	U' 2 $в/см^{1/2}$	P' 3 $вт/см^2$	R' 4 $ом\cdot см\cdot 10^6$	H 5 $ма/вт$	M, $г/см^2\cdot сек$ 6	$F(T)$	I'_e 7 $a/см^2$
2 300	1 319	0,1124	148,2	85,22	0,86	$7{,}8\cdot 10^{-10}$	0,352	0,041
2 400	1 422	0,1275	181,2	89,65	2,0	$4{,}3\cdot 10^{-10}$	0,375	0,116
2 500	1 526	0,1436	219,3	94,13	4,3	$2{,}03\cdot 10^{-9}$	0,399	0,298
2 600	1 632	0,1611	263,0	98,66	8,55	$8{,}4\cdot 10^{-9}$	0,422	0,716
2 700	1 741	0,1797	312,7	103,22	16,4	$3{,}2\cdot 10^{-8}$	0,446	1,631

1) I', amp/cm$^{3/2}$; 2) U', v/cm$^{1/2}$;
3) P', watt/cm^2; 4) R', ohm·cm·10^6;
5) H, ma/watt; 6) M, g/cm^2·sec;
7) I'_e, amp/cm^2.

It should be noted that the data for the evaporation of tungsten (the quantity M) are not accurate, since the rate of evaporation of tungsten to a large degree depends upon the purity of the metal and its structure.

The initial data for designing a tungsten cathode are normally the heating voltage U_n, the cathode emission I_e, and the service life (lifetime) τ of the cathode.

Determination of the operating temperature is the first step in designing a cathode. The following method may be employed to make a proper choice of the temperature. Equating the values for the cathode diameter from formulas (4-3a) and (4-8), we obtain:

$$\left(\frac{I_н}{I'}\right)^{\frac{2}{3}} = \frac{\tau M}{2{,}64}\cdot 10^4. \qquad (4\text{-}16)$$

For any temperature, the heating power may be determined as the ratio of the cathode emission to the efficiency at this temperature, $P_n = I_e/H$. Hence the heating current is $I_n = P_n/U_n = I_e/U_nH$. Substituting this value for the current into formula (4-16) and carrying through the bookkeeping, we obtain:

$$\frac{\tau^{3/2}U_н}{I_e} = \frac{4{,}29\cdot 10^{-6}}{I'HM^{\frac{3}{2}}}. \qquad (4\text{-}17)$$

All of the quantities in the denominator of the right side of this equation depend solely upon the cathode temperature. Letting $\Psi(T)$ stand for the right-hand side of the equation, we are able to find $\Psi(T)$ as a function of temperature, using the known parameters of the tungsten cathode. Figure 4-9 shows a graphical representation of the relationship between ln $\Psi(T)$ c and the operating temperature of the cathode.

Consequently, Eq. (4-17) may be rewritten as follows:

$$\frac{\tau^{\frac{3}{2}} U_{n}}{I_{e}} = \Psi(T). \tag{4-17a}$$

We find $\Psi(T)$ from the given values of heating voltage, emission current, and service life, and from the graph we find the value of temperature that corresponds to this value of Ψ.

In designing a real cathode, the initial emission-current value should be multiplied by the factor m, which takes into account the loss of emission owing to cooling at the ends. This factor depends upon the cathode temperature distribution, and differs for various

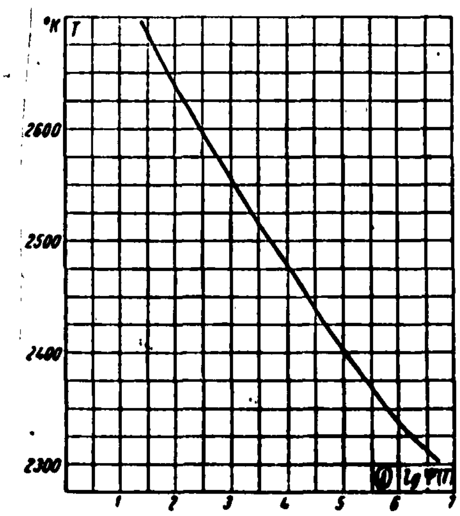

Fig. 4-9. Graph for choosing the operating temperature of a tungsten cathode.
1) log.

cathode designs, dimensions, and mounts. It is accurate enough to assume m = 1 + 0.15 n (n is the number of cooled ends).

Consequently, it is necessary to substitute the emission current mI_e into Eq. (4-17a).

For the cathode temperature found, the cathode diameter is determined in accordance with formula (4-3a).

$$d_к = 3{,}79 \cdot 10^3 \tau M. \qquad (4\text{-}18)$$

If τ expressed in hours, and M in g/cm^2·sec, the diameter turns out to be in centimeters.

Computing the cathode diameter, it is possible to find the magnitude of the heating current from formula (4-8).

The cathode length is determined from a formula derived from Eq. (4-12):

$$l_к = \frac{\sqrt{d_к}}{U'}(U_н + n\Delta U). \qquad (4\text{-}19)$$

In determining the cathode geometry, it is necessary to check the magnitude of the emission current. To do this, we compute the effective cathode length from Eq. (4-13). The calculated emission current is defined as the product of the specific emission at the operating temperature and the effective cathode surface:

$$I_{ep} = I'_e \pi d_к l_{эф}. \qquad (4\text{-}20)$$

The calculation is assumed to satisfy the initial conditions if $I_e \leqslant I_{ep}$.

4-5. FILM CATHODES. DESIGNING A CARBIDE-TREATED CATHODE.

The most common of the thermionic film cathodes are the thoriated carbide-treated tungsten cathodes. The feature of these cathodes is

the presence of a one-atom-thick film of thorium which, decreasing the work function, permits an increase in the thermionic emission.

The thoriated cathode. A thoriated cathode consists of a wire manufactured from powdered tungsten containing about 1.5 per cent thorium. An increase or decrease in the thorium oxide content is undesirable, since with an excess of thorium, the wire becomes brittle, and a low thorium content may prove inadequate to provide stable emission.

At a temperature of 2600°K, a thoriated cathode yields the same emission as does pure tungsten. This is explained by the fact that this temperature is not sufficient for the formation of an atomic film of thorium on the surface of the wire; to obtain such a layer, the cathode must be subjected to so-called activation. For this purpose, the cathode is baked in a vacuum at a temperature of 2800°K for 1-2 min. At this temperature, the thorium begins to be reduced, with the formation of pure thorium, which diffuses to the surface of the cathode, but which evaporates from it at this temperature. Next, the temperature is dropped to 2100-2200°K, and the cathode baked at this temperature for 15-20 min. With this heating, reduction of the thorium oxide ceases, but the previously formed thorium atoms continue to migrate to the surface of the cathode; thus, a greater number of thorium atoms appear on the cathode than evaporate from it. As a result, a monatomic film of thorium is formed on the cathode. A film several atomic layers thick cannot form, since at these temperatures tungsten strongly retains only those thorium atoms that are located directly on its surface.

The work function of a thoriated cathode when its surface is completely covered with a one-atom-thick layer of thorium is 2.6 ev, i.e. 2 ev less than the work function of pure tungsten.

Thoriated-tungsten cathodes operate at temperatures of 1800-1900°K. At these temperatures, their efficiency amounts to 25-40 ma/watt, while at the same time, the efficiency of pure tungsten cathodes working at temperatures of 2400-2700°K amounts to only 2-15 ma/watt. Thus, thoriated cathodes are at least four times as efficient as tungsten cathodes.

The service life of a thoriated cathode is determined by the time required for the thorium surface film to be destroyed, resulting in a decrease in the proportion of the surface covered by the film. This phenomenon is caused by the gradual depletion of the reserves of thorium in the tungsten owing to evaporation. When this happens, the number of thorium atoms diffusing to the cathode surface proves to be less than the number evaporating. In addition, diffusion of the thorium atoms to the surface is impeded by the growth of tungsten crystals at the working temperature of the cathode.

A serious drawback with a thoriated cathode, caused by the weak association of the thorium atoms with the tungsten surface, is its great sensitivity to overheating and ion bombardment. Under a brief thermal overload, only the surface film of thorium evaporates, and cathode emission may be restored by reactivation. With extended overheating, both the surface thorium and the deep thorium are driven off, and the cathode cannot be restored.

Destruction of the thorium film by ions of residual gases occurs when there are electrical fields in the tube that are strong enough to accelerate the ions to considerable velocities. At plate voltages above 300 v, cathode destruction owing to ion bombardment becomes appreciable. These thoriated-cathode faults limit the applications of this cathode. Almost no thoriated cathodes are utilized in electron tubes.

The carbide-treated cathode. Studies of the thoriated cathode have shown that its life is increased and emission becomes more stable if the thoriated tungsten is carbide-treated, i.e., part of the tungsten is converted into tungsten carbide (W_2C) by heating the cathode before it is activated in vapors of any hydrocarbons (naphthalene, gasoline, etc.) at a temperature of 1600°K. When this is done, the vapors in contact with the surface of the hot tungsten dissociate. The carbon forming combines with the tungsten. As a result, a tungsten carbide jacket forms on the cathode. With complete carbide-formation, in which all of the tungsten is converted into tungsten carbide, the cathode becomes very brittle and unfit for working. Thus, the carbide-treatment process is normally halted when the carbon content of the tungsten reaches 0.2-0.5 per cent (the tungsten is sheathed with tungsten carbide), and then the material is activated, in a manner similar to the way in which a thoriated cathode is activated. It is easier to activate a carbide-type cathode, however, since the thorium is reduced from the oxide not just by the tungsten, but chiefly by the carbon:

$$ThO_2 + 2C = Th + 2CO.$$

The atoms of thorium appearing in the process of cathode activation migrate to the surface and form a monatomic film, which is located not upon the pure tungsten, as in the case of a simple thoriated cathode, but on the tungsten carbide.

The active thorium film is more strongly held on the tungsten carbide surface than on the surface of pure tungsten. Thus, at precisely the same temperature, the evaporation of thorium takes place several times more slowly from the tungsten carbide surface than from a pure tungsten surface, which permits a certain increase in the

working temperature of the carbide cathode (up to 1950-2000°K), and a rise in efficiency up to 50-70 ma/watt with a lifetime of from 1000 to 3000 hr. In addition, for the same reason, the active layer of thorium in a carbide-type cathode is less sensitive to overheating and to the effect of high plate voltage. These advantages make it possible to utilize a carbide- type cathode in high-power oscillator tubes with plate voltages of several kilovolts, and in low- and medium-power oscillator tubes.

Designing a thoriated carbide cathode. In comparison with pure tungsten, the resistance of the wire and the radiating ability of a thoriated carbide-type cathode are greatly decreased, owing to carburization and the formation of tungsten carbide.

The change in these quantities depends upon the carbon content of the tungsten, which is difficult to determine with any great accuracy. As a rule, a carbide-type cathode is designed for a temperature of 2000°K; it is assumed that the tungsten carbide occupies 30 per cent of the cross-sectional area of the cathode wire. This is the optimum ratio of tungsten carbide to pure tungsten. With this ratio and a temperature of 2000°K it is possible to determine all of the quantities needed for design purposes.

The efficiency of a carbide-type cathode at a temperature of 2000°K may be taken as 55-65 ma/watt. From this, and the given value of cathode emission current, the cathode heating power and current are found.

Determining the cathode heating current, we may find its diameter from formula (4-8). The heating current for a unit carbide-type cathode at a temperature of 2000°K is 1022 amp/cm$^{3/2}$. The cathode diameter is found from the following relationship:

$$d_к = 0{,}99 \cdot 10^{-2} I_н^{2/3} [см]. \quad (4\text{-}21)$$

The resistance of a unit carbide-type cathode at 2000°K is $R'_k = 88 \cdot 10^{-6}$ ohm·cm. Since the cathode resistance is $R = R'_k l_k / d_k^2$, setting it equal to the Ohm's-law resistance, we obtain:

$$\frac{U_н}{I_н} = R'_к \frac{l_к}{d_к^2}. \quad (4\text{-}22)$$

Where the Ohm's law resistance is $R = U_n/I_n$.

In designing a real cathode, the channel voltage should be increased by a correction allowing for the drop in cathode resistance due to cooling at the supported ends. At a temperature of 2000°K, the correction for each cooled end is ΔU = 0.21 v. Rearranging Eq. (4-21) and substituting in the value of R'_k, we obtain an expression for the length of the cathode:

$$l_к = 1{,}13 \cdot 10^4 \frac{U_н + n\Delta U}{I_н} d_к^2 [см]. \quad (4\text{-}23)$$

In this expression, the voltage is in volts, the heating current in amperes, and the cathode diameter in centimeters.

4-6. OXIDE CATHODES.

The basic type of cathode, which has large thermionic emission at comparatively low temperatures, is the so-called oxide cathode, which takes the form of a metal base to which is applied a layer consisting of a mixture of oxides of the alkali metals Ba, Sr, and Ca with a small amount of pure barium (Fig. 4-10). It has been established that in order to obtain any considerable thermionic emission, it is necessary to have pure barium both within and on the surface of the oxide layer.

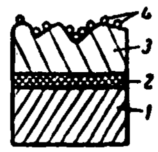

Fig. 4-10. Structure of
an oxide-coated cathode.

The high thermionic emission of the oxide-coated cathode was first observed in 1904 in an investigation of the oxides of barium, strontium, and calcium, applied to platinum and subjected to appropriate heat and electrical treatment. At the present time, oxide-coated cathodes are manufactured in the following manner. The metal base, also called the core or base layer, is covered with a layer of carbonates, alkali-earth metals (carbonates) $BaCO_3$, $SrCO_3$, and $CaCO_3$. Modern oxide-coated cathodes are covered with either a triple carbonate consisting of a mixture of barium, strontium, and calcium carbonates in various per cent ratios, or with a double carbonate consisting of a mixture of barium and strontium carbonates. Simple carbonates, for example, barium carbonates, are not utilized. For purposes of application to the metal base layer, suspensions of the given carbonates are prepared in an organic liquid, or in water with the addition of a binding material (binder). Several methods are in use for applying the suspension to the metal base layer. The simplest method, but a method that is already obsolete, is to spread the material on; this method is utilized only in isolated cases. A more modern method is the method of dipping or of pulling the core through the suspension. These two methods are utilized chiefly in manufacturing directly heated filaments. In applying a suspension to indirectly heated cathodes, it is preferable to utilize the so-called method of atomization (spraying), which makes it possible to obtain a layer that is sufficiently uniform throughout its thickness.

Finally, there is a last method for applying a suspension, the method in which the carbonates are applied to the base layer by the motion of charged colloidal particles of the carbonate in a constant electric field, and their deposition on the base layer (cataphoresis). This method makes it possible to obtain very thin (to 0.5 micron) dense coatings.

After application of the suspension, the cathode is mounted in the device and subjected to appropriate heat treatment in vacuum, with continuous exhausting of the evolving gases. In order to break the carbonate down to the oxides, the cathode is heated to a temperature of 1500-1600°K. The heating-up process is carried out gradually, since with a sharp change in temperature, owing to nonuniform heating up of the core and the oxide layer, the oxide layer will flake off. At temperatures of 1500-1600°K, the carbonates are transformed into oxides of the alkali-earth metals accompanied by liberation of carbon dioxide, which is exhausted from the device:

$$BaCO_3 \rightarrow BaO + CO_2.$$

The oxide layer thus obtained yields almost no noticeable thermionic emission, since within the layer, and on its surface, there are no free barium atoms. Thus, further treatment of the cathode is designed to partially reduce the barium from its oxide. This is done by heating the cathode, using a temperature of 1200°K, and simultaneously causing emission current to flow, using a positive voltage on the plate on the order of 100-200 v. When this is done, the barium is reduced both as a result of chemical interaction with active impurities in the base material, and owing to electrolysis of barium oxide owing to the passage of the emission current through the oxide layer.

There are always chemically active impurities such as, for

example, silicon, magnesium, and aluminum present in the grades of nickel used for a base. Thus, at temperatures of 1200°K, chemical reactions take place with the formation of pure barium in the oxide layer. Of course, when this is done, oxides of the SiO_2 type form at the base-layer — oxide-layer interface, which noticeably increases the resistance facing the emission current. This effect is particularly substantial in cathodes operating in pulse mode with large emission currents being taken off the cathodes. Thus, for oxide-coated cathodes in pulsed operation, nickel of high chemical purity is utilized.

In addition, barium reduction may occur as a result of the interaction of oxides with the carbon monoxide liberated in the layer when the binder is burned off.

The positive barium ions that form upon electrolysis of the oxides migrate to the surface of the base layer under the influence of the electric field; here they are neutralized, and then diffused throughout the oxide layer, and to its surface. Reduction of barium owing to electrolysis of barium oxide occurs through a reaction of the $BaO \rightleftarrows \overset{++}{Ba} + \overset{--}{O}$ type. The oxygen that evolves is either exhausted from the device, or absorbed by a getter. After the cathode has been activated, it is subjected to so-called aging, in which it is placed under comparatively large-load conditions for a relatively long period of time. During the aging process, there is additional activation of the cathode by the plate current, and equalization of the distribution of barium atoms with respect to the volume and surface of the oxide layer takes place. As a result of aging, cathode operation becomes more stable. This is also facilitated by the additional outgassing of other electrodes due to electron bombardment during aging. In this case, the thermionic emission first drops somewhat owing to poisoning of the cathode by gases evolving from the other electrodes, but later the

emission again increases, and by the end of aging reaches a stable value.

The base-layer material proves to have a strong influence on the emission of an oxide-coated cathode. The base layer of an oxide-coated cathode should be stable at the operating temperature, and at the temperature utilized in treating the cathode. This requirement is met by metals having high melting points and low rates of dispersion. In addition, the base layer of an oxide-coated cathode should have a favorable effect upon cathode emission. This requirement may be met by metals that contain chemically active impurities and admixtures that aid activation of the cathode. Finally, heat treatment of the cathode should result in good outgassing of the base-layer material.

At the present time, bases for oxide-coated cathodes are manufactured almost exclusively from nickel and tungsten. Nickel is utilized for the base layers of indirectly heated oxide cathodes and coated filaments in the form of relatively thin wires and ribbons. The base layers of thin oxide-coated filaments are normally made of tungsten, sometimes first coated with a thin layer of an alloy of copper with 6 per cent aluminum (aluminum bronze). The presence of the thin layer of bronze on the surface of the tungsten base layer protects the tungsten against oxidation when it is heated in a furnace with the applied suspension of carbonates. With the cataphoretic method for applying carbonates, it is not necessary to bronze the base layer.

The low work function of an oxide cathode (about 1.1 ev) makes it possible to obtain large thermionic emission (0.3-0.5 amp/cm^2) at comparatively low operating temperatures (900-1150°K). Here the efficiency of well-activated oxide cathodes amounts to 60-100 ma/watt, and the lifetime to 1000-1500 hr and more. The life of an oxide cathode depends basically upon three factors: the formation of a barrier layer

at the base-metal — oxide interface, poisoning of the cathode, and destruction of the oxide layer.

Upon extended service of the cathode, there may form at the base-material — oxide coating interface an intermediate layer of a chemical compound of oxides with active impurities of the base material: silicon, aluminum, titanium, magnesium, etc. This compound may have the form, for example, of Ba_2SiO_4. Such a layer has very high resistance and, consequently, causes a deterioration in the conditions for taking current off from the cathode. The intermediate layer has the nature of a semi-conductor, i.e., its resistance rises as the temperature drops. Thus, the intermediate layer makes itself felt especially strongly when the cathode is operated underheated.

The mechanism for the formation of a barrier layer has not been sufficiently studied; experiments show, however, that it will grow more rapidly in cathodes operating at high temperatures.

Poisoning of an oxide-coated cathode is connected with the appearance on the cathode during the process of its operation of materials combining with free barium. The substances may appear owing to liberation of gases from the parts of devices when they are heated. The most dangerous gases for an oxide-coated cathode are oxygen and the halogens. The dusting of certain metals on the cathode, in particular gold and free copper, which form a solid solution with barium, may also lead to poisoning.

Destruction of the oxide layer occurs owing to nonuniformity in the emission properties of the cathode surface. Where a considerable amount of current is drained from the cathode, its entire surface is not involved, but only the well activated portion. The great current passing through these sections causes local overheating of the layer, which may lead to its destruction. This effect occurs especially

vigorously at a lower cathode temperature, where the resistance of the oxide layer is high and, consequently, a great deal of power is involved when the current is passed through the layer.

A serious drawback to an oxide cathode is the slow, but constant evaporation of barium and barium oxide from the surface of the oxide layer, which takes place at the cathode working temperatures. The barium and barium oxide evaporating from the surface of the oxide layer are deposited upon other electrodes of the tube, and decrease their work function. This results in the occurrence of thermionic and secondary-electron emission from these electrodes, and changes the magnitude of the contact potential differences between them, causing a change in the parameters and characteristics of the tubes.

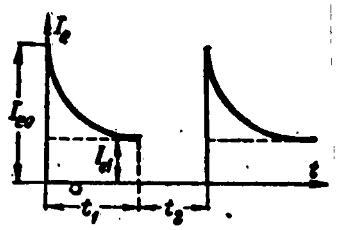

Fig. 4-11. Graph showing the change in the emission current of an oxide cathode with time where the plate voltage is applied briefly.

The faults noted somewhat limit the region of application of the oxide-coated cathode. For example, in its common form, the oxide-coated cathode cannot be utilized in devices operating at high plate voltages. They are successfully utilized, however, in receiving and amplifying tubes, in cathode ray tubes, in pulse generator tubes, and SHF devices.

Pulsed operation of an oxide-coated cathode. If we follow the change in emission current of an oxide-coated cathode in time, it turns out that the maximum value occurs during the first instants

that the plate voltage is applied (I_{e0} on Fig. 4-11), and it then sharply decreases to a certain constant value I_{e1}, which remains almost unchanged with time. This is explained by the fact that at the instant the plate current is applied, there are a large number of barium atoms on the surface of the oxide layer, and they provide a large emission current. As the emission current passes, the barium atoms lose their electrons, and are converted to positive ions, which migrate in the oxide layer toward the base layer of the cathode under the action of the electric field. After a time t_1, amounting to roughly 250-300 μsec, there is an amount of barium atoms remaining on the surface of the oxide layer which does not decrease further with time, since an equilibrium condition has been established, in which the number of positive ions of barium migrating to the base layer equals the number of neutral barium atoms diffusing to the surface of the oxide layer. A decrease in the number of barium atoms on the surface of the oxide layer leads to an increase in the work function and, consequently, to a decrease in emission to a value I_{e1}. Thus, if the plate has voltage applied to it for an extended period of time, an emission current of not more than I_{e1} may be taken from the cathode. If the plate voltage is disconnected for a short time t_2 amounting to about 0.001-0.01 sec, the surface of the oxide layer is enriched with barium atoms rather rapidly. As a result, the work function again drops. If we now connect in the plate voltage, the pattern is repeated: at the instant of application of the voltage, the emission current reaches a maximum value I_{e0}, which then rapidly drops to a value I_{e1}.

It has been established that when emission current is taken briefly from an oxide-coated cathode, the current density may reach 30-50 amp/cm^2 while when emission current is continuously withdrawn, the density all-in-all may reach 0.3-0.5 amp/cm^2. The ability of an

oxide-coated cathode to yield very high emission during brief time

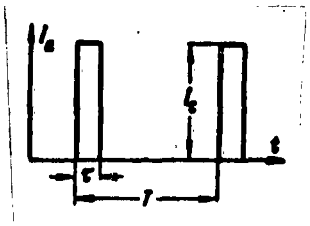

Fig. 4-12. Graph of the variation of plate current of an oscillator tube in pulsed operation.

intervals, and to regain emission properties during brief nonworking intervals has been successfully utilized in tubes working in pulse mode. For example, in pulse oscillator tubes, the operating time intervals of length τ alternate with nonworking intervals of length T (Fig. 4-12). Here, during the pulse length τ, the cathode provides a pulse of several tens of amperes of emission current, with an average emission current amounting to several tens of milliamperes. Such large emission in pulsed operation is obtained with small geometric dimensions, and with low heating power of the oxide-coated cathode.

Surface irregularities in the oxide layer have a strong effect upon the pulse emission of an oxide-coated cathode. As a rule, the maximum values of pulse emission are obtained with rough oxide-layer surfaces.

4-7. DESIGNING OXIDE-COATED CATHODES.

Designing an oxide-coated cathode amounts to determining its geometric dimensions and heating power, and also (in the case of indirectly heated cathodes) the geometric dimensions of the heater as well.

The basic data for designing cathodes are the heating voltage, rated value of operating current from the cathode, and, sometimes, the peak value of the current taken from the cathode. In certain types

of tubes, the cathode geometry is affected by other given parameters of the tube. In this case, design of the cathode amounts to determining the heating current and voltage, and the heater dimension.

a) Designing indirectly heated cathodes. The design of indirectly heated cathodes begins with the choice of the maximum cathode temperature (the temperature of the cathode at the hottest spot), and the magnitude of the operating current density of the cathode. The choice of these quantities is connected with the design of the device, and the life set for the device. Sample values for these quantities, for various types of devices, are given in Table 4-2.

TABLE 4-2.

1) Вид приборов	2) Максимальная температура $T_м$, °K	3) Плотность тока с катода J_p, ма/см²
4) Стеклянные и металлические приемно-усилительные лампы	950—1 100	16—20
5) Выпрямительные лампы и выходные лампы (для усиления мощности)	1 050—1 150	30—90
6) Пальчиковые и миниатюрные лампы	1 000—1 100	30—60
7) Генераторные лампы	1 050—1 150	25—80
8) Импульсные приборы с оксидным катодом	1 100—1 150	5 000—10 000

1) Type of devices; 2) maximum temperature T_M, °K; 3) density of current taken from cathode, j_p, ma/cm^2; 4) glass and metal receiving-amplifying tubes; 5) rectifier tubes and output tubes (for power amplification); 6) button-type and miniature tubes; 7) oscillator tubes; 8) pulse devices with oxide-coated cathodes.

The choice of cathode working temperature greatly depends upon the structure and parameters of the device in which the cathode is to be used. From the point of view of lifetime and operating stability of the devices, it is desirable to use cathodes having the maximum permissible temperature. For a cathode to operate well at the minimum

permissible temperature, however, it is necessary to carry out careful production treatment of both the cathode itself and the other parts of the device. This does not always turn out to be possible.

In choosing the working temperature of a cathode, it is also necessary to take into account the temperature of the parts of the device surrounding the cathode, which cause additional heating up of the cathode.

For the value of operating cathode current density j_p chosen, it is possible to determine the size of the oxide-coating surface:

$$F_{\text{окс}} = \frac{I_\text{к}}{j_\text{р}},^{*} \tag{4-24}$$

where I_k is the nominal current drawn from the cathode.

The cathode is designed in accordance with the value found for the surface of the oxide-coated portion of the cathode. In the case of tubular cathodes, examples of which are shown in Fig. 4-5, it is necessary to determine the tube diameter, length of the sections not oxide-coated, the distance from the end of the oxide coating to the insulators. The length of the noncoated sections is so chosen that

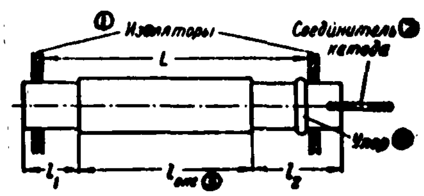

Fig. 4-13. Construction of a tube for an indirectly heated cathode.
1) Insulator; 2) cathode connection; 3) l_{oks}; 4) support.

the gap between the oxide coating and the insulators, where the cathode is mounted, will be not less than 0.8-1.2 mm. In addition,

---

* [F_{okc} - F_{oks} - $F_{oksidnyy}$ - F_{oxide}.]

it is necessary to take into account the need to extend a portion of the tube beyond the insulator with a length adequate for welding on the cathode connection. Figure 4-13 shows the structural dimensions of a tubular cathode. A support (collar) is normally placed on one end of the cathode; it keeps the cathode from moving axially through the hole in the insulator.

The length of the oxide coating is found according to the chosen diameter of the tube. For a cylindrical tube

$$l_{окс} = \frac{F_{окс}}{\pi (d_к + 2\Delta_{окс})}, \qquad (4\text{-}25)$$

where Δ_{oks} is the thickness of the oxide-coating layer, which may vary in magnitude from 0.05 to 0.1 mm.

For a tube of any cross-sectional shape, whose perimeter is p:

$$l_{окс} = \frac{F_{окс}}{p}. \qquad (4\text{-}26)$$

In choosing the diameter of the tube, it is desirable to keep the ratio of the distance between the insulators to the tube diameter within values of 8 to 20.

With very short cathodes, there is a strong increase in cooling by the insulators and cathode connection, which increases the consumption of power for heating. On the other hand, where the diameter of the cathode is large relative to the length, its rigidity drops, which may lead to bending of the cathode during tube assembly and upon thermal expansion of the cathode.

Calculation of the heating parameters of a cathode reduces to the determination of the specific heating power required to provide the required cathode temperature. For an ideal cathode, the specific heating power equals the power radiated from unit surface of the oxide layer at the given temperature. Figure 4-14 shows graphically the

dependence of the power radiated from 1 cm^2 of oxide surface upon the

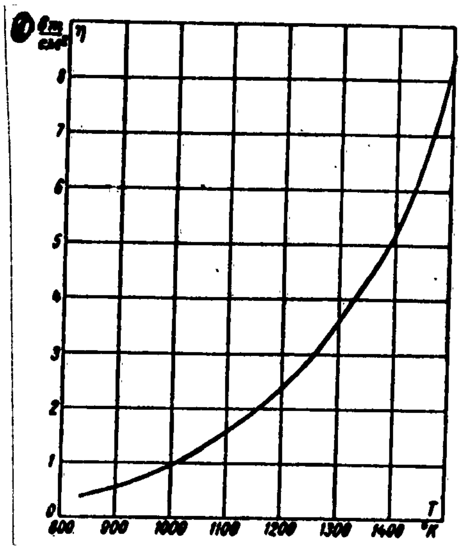

Fig. 4-14. Radiant power of an oxide-coated cathode as a function of temperature.
1) watt/cm^2.

temperature; the graph is based upon results of measurements performed by Yu. F. Zarutskiy.

The heat carried off by insulators and the base eyelet are of substantial importance in the thermal balance of a heated cathode. As a result of this cooling, there is a drop in cathode temperature both in the hottest spot, and at the ends of the cathode. The drop in temperature at the ends of the cathode is greater, and therefore the temperature distribution over the cathode is nonuniform. Figure 4-15 shows the temperature distribution over the cathode of a 6N9S tube.

The drop in temperature of the ends of a cathode depends not only on the heat carried off by insulators and the cathode connection, but also to a certain degree upon losses owing to radiation from the ends of the cathode tube. Thus, it is possible to effect a certain decrease in the temperature nonuniformity, on occasion, by flattening the ends of the tube above the insulator (Fig. 4-16).

The proportion of heat carried off owing to cooling with respect to the total power consumed in heating is the greater the smaller the distance between the insulators in which the cathode is mounted.

As the length of the cathode increases, there is a rise in the power radiated from its surface, while the power carried off by the insulators remains unchanged. Consequently, the influence of cooling drops as the cathode is made longer.

A certain drop in the power carried off by the insulators may be obtained by decreasing the cathode-tube – insulator contact surface. To do this, the openings for the cathode in the insulators are frequently made with additional slots, as shown in Fig. 4-17.

As a result of changing the amount of heating power used in heat that is carried off, when the length of the cathode is changed, the specific heating power required to obtain the necessary maximum temperature depends mainly upon the separation of the insulators. Figure 4-18 gives curves showing the specific heating power as a function of cathode length for various values of maximum temperature.

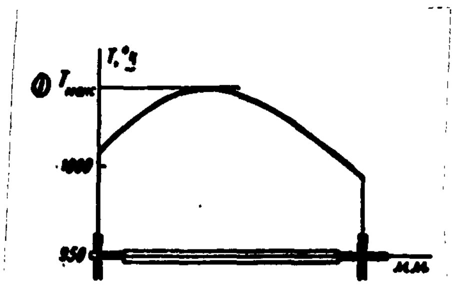

Fig. 4-15. Temperature distribution over the cathode of a 6N9S tube. 1) T_{max}.

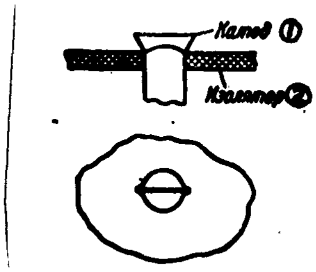

Fig. 4-16. Flattenin the end of a cathode tube. 1) Cathode; 2) insulator.

The curves given correspond to cathodes with a 0.05-mm thick nickel tube and 0.4-mm thick mica insulators. Variation of these dimensions within normally permissible limits – cathode material up to 0.1 mm thick and mica insulators 0.3-0.5 mm thick – has practically no effect

upon the value of the maximum cathode temperature. Thus, these curves

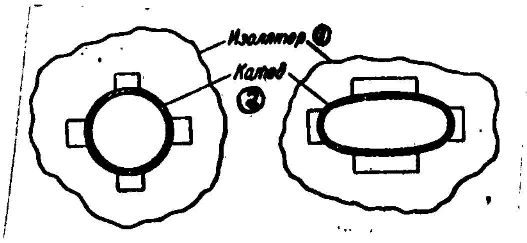

Fig. 4-17. Shape of cathode openings in insulator designed to decrease heat losses.
1) Insulator; 2) cathode.

are suitable for obtaining the specific heating power for a majority of heater-type cathode designs in use.

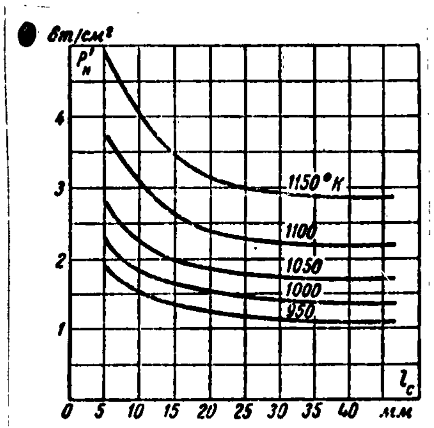

Fig. 4-18. Curves for specific heating power as a function of insulator separation.
1) watt/cm^2.

Using the curves showing specific heating power as a function of insulator separation, it is possible in designing a cathode to find the required specific heating power. The total heating power will then be found as the product of the specific power and the total area of the side surface of the cathode F_k:

$$P_н = P'_н F_к. \qquad (4\text{-}27)$$

For a given heating voltage and a known heating power, the heating current is found:

$$I_н = \frac{P_н}{U_н}. \qquad (4\text{-}28)$$

After calculation of the heating parameters, the heater is designed.*

b) Heater design. The most common type of heater for an oxide-coated cathode is a heater made of a tungsten wire (or of tungsten alloy), coated with an alundum insulating layer. The thickness of the alundum coating is so chosen as to provide adequate electric insulating strength in the presence of potential differences between cathode and heater not exceeding, for the majority of devices, 100-150 v, as well as mechanical strength for the layer. For common types of heaters, the thickness of the alundum coating is chosen so as to lie within 60-80 microns.

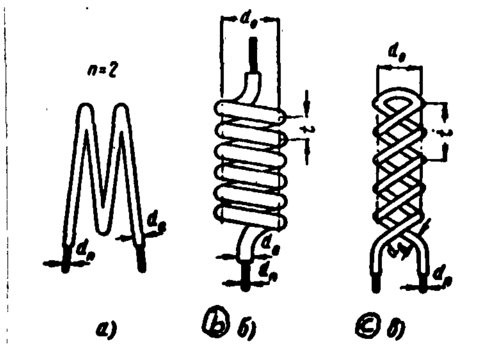

Fig. 4-19. Designations for heater dimensions.

The heater takes the form of either a single loop or several loops, or else is wound as a simple helix heater or double-helical heater (Fig. 4-19). Designing such a heater for given heating

* The system suggested for designing a heater-type cathode and heater is borrowed from Yu. F. Zarutskiy.

parameters and cathode geometry reduces to determining the diameter and length of the heater wire.

The resistance of a heater at the operating temperature, determined from known values of heating current and voltage, is connected with the diameter and length of the wire of the heater by the well-known expression

$$R_n = \frac{U_n}{I_n} = \rho_t \frac{4l_n}{\pi d_n^2}.^{*} \qquad (4\text{-}29)$$

It is clear that the problem of choosing heater geometry has no one single solution, since it is possible to pick an infinite number of combinations of the quantities l_p and d_p that satisfy Eq. (4-29). For a given heater length, the diameter of the wire may be determined uniquely:

$$d_n = \sqrt{\frac{4\rho_t l I_n}{\pi U_n}}. \qquad (4\text{-}30)$$

Thus, to design a heater it is necessary to make a proper choice of length and to determine, albeit roughly, the temperature; then we may find the magnitude of the resistivity ρ_t.

For a loop-type heater, the length of the loop should equal the length of the cathode tube. The total length of a heater (without the free ends used for connecting in the current-carrying electrodes) equals:

$$l_n = 2nl_k. \qquad (4\text{-}31)$$

where n is the number of loops, and l_k is the total length of the cathode tube. The number of loops is found from structural considerations designed to insure that the heater may be freely inserted in the

---

* [R_n - R_p - $R_{podogrevatel'}$ - R_{heater}.]

cathode tube. In an approximate computation, the number of loops may be found from the empirical formula

$$n \approx (2 \div 3) S_{к}, \tag{4-32}$$

where S_k is the cross-sectional area of the cathode tube, mm^2 (n should be an integer).

The total length of a helix heater is found from the expression

$$l_{н} = \frac{\sqrt{\pi^2 d_0^2 + t^2}}{t} l_{к} \tag{4-33}$$

while for a double-helical heater it is

$$l_{н} = \frac{2\sqrt{\pi^2 d_0^2 + t^2}}{t} l_{к} \tag{4-34}$$

(with the designations for the quantities as shown in Fig. 4-19). The quantity d_0 may range from 0.5 to 0.7 d_{k0}; d_{k0} is the inside diameter of the cathode.

Determining the length of the heater wire, we must find its temperature. The temperature calculation is based upon the fact that heat transfer from a heater to the cathode occurs solely owing to radiation of the heater surface. Heat transfer owing to direct contact of the heater surface with the interior surface of the cathode tube is negligible in practice, and may be neglected in the calculation.

Consequently, heat transfer between the cathode and heater obeys the Stefan-Boltzmann equation for two heated bodies with different radiation coefficients. Under steady-state conditions, i.e., with the temperatures of cathode and heater constant, all of the heating power is radiated by the heater surface:

$$P_{н} = S_{н\phi} \varepsilon_{пр} \sigma (T_{н}^4 - T_{к}^4),^* \tag{4-35}$$

* [$\varepsilon_{пр}$ - ε_{pr} - $\varepsilon_{privedenyy}$ - $\varepsilon_{reduced}$.]

where S_{eff} is the effective heater surface; T_p and T_k are the temperatures of the heater and the cathode; ε_{pr} is the reduced radiation coefficient for the cathode-heater system, which may be determined from the formula

$$\frac{1}{\varepsilon_{np}}=\frac{1}{\varepsilon_a}+\frac{p_{э\phi}}{p_\kappa}\left(\frac{1}{\varepsilon_\kappa}-1\right). \qquad (4\text{-}36)$$

In this formula, ε_a is the integral coefficient for radiation from the surface of the alundum-treated heater, which has in the working-temperature interval a value of 0.18-0.22. Where the alundum coating is very thin (less than 40 microns), this quantity increases to 0.24-0.28, owing to the fact that the alundum layer is transparent with respect to radiation directly from the base layer of the heater. The integral radiation coefficient for the interior surface of the cathode $\varepsilon_k = 0.17$ for nickel. The quantities p_{eff} and p_k are respectively the effective heater perimeter and the interior perimeter of the cathode tube.

Generally speaking, the effective perimeter of the heater does not equal the sum of the perimeters of the individual filaments (for a loop-type heater) since the filaments shield each other from the cathode. The degree of shielding depends upon how well the heater fills the interior volume of the cathode, and upon the configuration of the cathode tube. For a cylindrical cathode with a loop-type heater, a theoretical computation has been made for the quantity $k = p_{eff}/p_k$ as a function of the ratio of the heater-filament diameter to the inside diameter of the cathode tube. Figure 4-20 shows curves of this function for various numbers of heater loops. On the basis of the value found for k, it is possible to determine the effective perimeter:

$$p_{э\phi}=kp_\kappa. \qquad (4\text{-}37)$$

In this case, the effective surface is found from the expression

$$S_{эф} = p_{эф} l_к. \tag{4-38}$$

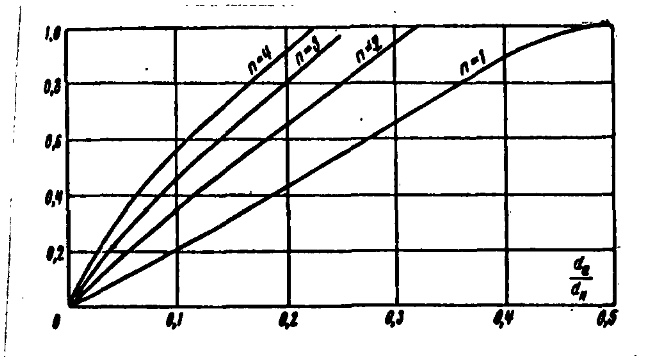

Fig. 4-20. Chart for determining surface fill factor.

With a helix or double-helical heater, it is possible to find the effective surface by multiplying the total surface of the spiraled filament by the shielding factor δ:

$$S_{эф} = \frac{\pi d_a l_к}{t} \delta \sqrt{\pi^2 d_0^2 + t^2} \tag{4-39}$$

for a helix and

$$S_{эф} = \frac{2\pi d_a l_к}{t} \delta \sqrt{\pi^2 d_0^2 + t^2} \tag{4-40}$$

for a double-helical heater.

The shielding factor δ is a function of the spiral fill factor α. For helix heaters, $\alpha = d_a/t$, while for double-helical heaters, $\alpha = 2d_a/t$. Figure 4-21 gives a graph of $\delta = f(\alpha)$. The effective perimeter of such heaters may be found from the expression

$$p_{эф} = \frac{S_{эф}}{l_к}. \tag{4-41}$$

From this may be determined the coefficient $k = p_{eff}/p_k$.

Thus, for any heater design, it is possible to express the

effective surface in terms of the cathode perimeter, cathode length, and the coefficient k:

$$S_{эф} = k p_к l_к. \qquad (4\text{-}42)$$

Substituting the value of the effective heater surface into the Stefan-Boltzmann equation, and the value of the reduced radiation coefficient, and expressing the heater temperature in terms of this equation, we obtain:

$$T_п^4 = T_к^4 + \frac{P_н}{l_к p_к \sigma}\left(\frac{1}{\varepsilon_п k} + \frac{1}{\varepsilon_к} - 1\right). \qquad (4\text{-}43)$$

The basic quantities, in practice, for designing a heater are d_p and, consequently the quantity d_a associated with it. Thus, the coefficient k in expression (4-43) is an unknown. It is clear that its maximum possible value is 1, which corresponds to complete filling of the space within the cathode by the heater. Analysis of expression (4-43) shows that the maximum value of the coefficient k corresponds to the lowest possible heater temperature, and equals:

$$T_{п.мин}^4 = T_к^4 + \frac{P_н}{l_к p_к \sigma}\left(\frac{1}{\varepsilon_п} + \frac{1}{\varepsilon_к} - 1\right).^* \qquad (4\text{-}44)$$

In first approximation, in calculating the heater temperature, we must assume that k = 1, and find the minimum hea. temperature. From this temperature it is possible to find the resistivity ρ_t (Fig. 4-22) and from expression (4-30) it is possible to determine the heater-wire diameter d'_p, corresponding to the maximum value of the surface fill factor. According to the quantity d'_p obtained and the corresponding value d'_a, we find the actual value of the coefficient k, in first approximation. Using this value, for a second approximation,

---

* [мин - min - minimum.]

the temperature is again found from Eq. (4-43), in accordance with which the quantity d_p is recalculated. As a rule, the second approximation gives a result that is accurate enough for practical purposes.

In designing a heater, it is necessary to add to the calculated length of the wire the length of the ends needed to make the connection to the current-carrying electrodes.

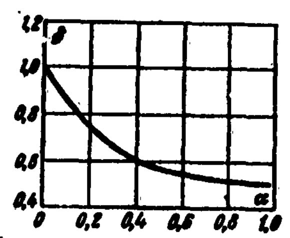

Fig. 4-21. Chart for determining the screening factor for helix and double-helical heaters.

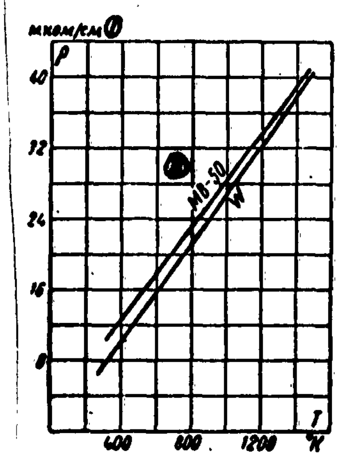

Fig. 4-22. Graph of resistivity of tungsten and MV-50 alloy as a function of temperature.
1) ohm/cm; 2) MV-50.

In many cases, it turns out to be necessary to recalculate the length or diameter of the wire according to the known parameters for heating of a real heater or to replace the base-layer material. In such a case, it is possible to utilize the formula suggested by V.F. Kovalenko:

$$1{,}29\frac{\Delta l_n}{l_n} - 0{,}71\frac{\Delta U_n}{U_n} - 2{,}18\frac{\Delta d_n}{d_n} + 0{,}82\frac{\Delta l_n}{l_n} + \frac{\Delta \rho}{\rho} = 0. \qquad (4\text{-}45)$$

Each term of this formula contains a ratio of the required change in one of the quantities characterizing the heater (heating current and voltage, wire diameter and length, and ρ) to the same quantity. If any of these quantities remains unchanged, the appropriate term of the

equation will become zero. Thus, for example, where the length of the wire is recalculated in order to provide a necessary change in heating current while maintaining the other quantities constant, Eq. (4-45) takes the following form:

$$1{,}29\frac{\Delta I_n}{I_n}+0{,}82\frac{\Delta l_n}{l_n}=0. \qquad (4\text{-}45a)$$

In like manner, where the diameter is changed (for example, in order to utilize a standard wire diameter),

$$0{,}82\frac{\Delta l_n}{l_n}-2{,}18\frac{\Delta d_n}{l_n}=0 \qquad (4\text{-}45b)$$

while where the heater material is replaced (changing the value of ρ), with the heating parameters unchanged

$$\frac{\Delta\rho}{\rho}-2{,}18\frac{\Delta d_n}{d_n}+0{,}82\frac{\Delta l_n}{l_n}=0. \qquad (4\text{-}45c)$$

c) Designing filamentary cathodes. For standard filaments, the thickness of the oxide coating is on the order of the base-material diameter. As a result, the temperature of the filament base material may be considerably different from the surface temperature. According to some data, the temperature drop may reach 40-50°. It is especially important to allow for the difference in the temperatures of the surface of the cathode and of the base material with very thin bases (with diameters on the order of 10-30 microns). For bases of larger diameter, and where the oxide coating is not very thick ($d_{oks}/d_k \leqslant 3$) and in an approximate calculation for a cathode, the temperature drop may be neglected.

The design of a cathode begins with the choice of the working temperature and efficiency, average values of which are given in Table 4-3.

TABLE 4-3.

Вид приборов (1)	(2) Максимальная рабочая температура керна $T_к$, °K	(3) Средняя рабочая эффективность H_p, ма/вт
(4) Стеклянные приемно-усилительные лампы	950—1 050	20—40
(5) Пальчиковые и миниатюрные приемно-усилительные лампы	1 000—1 250	30—70
(6) Выпрямительные лампы	1 050—1 150	10—30
(7) Генераторные лампы и лампы для усиления мощности	1 050—1 200	30—80

1) Type of devices; 2) maximum operating temperature of base, T_K, °K; 3) mean operating efficiency N_p, ma/watt; 4) glass receiving-amplifying tube; 5) button-type and miniature receiving-amplifier tubes; 6) rectifier tubes; 7) oscillator tubes and power-amplifier tubes.

The upper temperature limits refer to cathodes with a base consisting of a thin wire. Here the temperature of the oxide-layer surface proves to be somewhat less.

The basic data for designing directly heated cathodes (as in the design of an indirectly heated cathode) are the heating voltage and the magnitude of the emission current drawn under working conditions.

On the basis of the chosen value of efficiency, the cathode heating power required is determined:

$$P_н = \frac{I_з}{H_р} \quad (4\text{-}46)$$

and the heating current is

$$I_н = \frac{P_н}{U_н}.$$

In certain tubes, for example, in rectifiers, rather than working with the mean value of cathode current, the peak value is used, which is several times the mean-current value. In this case, the heating

power should be defined as the ratio of the peak cathode current to the maximum possible cathode efficiency, equal for filamentary cathodes to 100-120 ma/watt.

From the heating parameters found, the cathode geometry is calculated. The following formulas are used in calculations for large-diameter cathodes:

$$P_{\text{н}} = \eta_{\text{окс}} \pi d_{\text{окс}} l_{\text{н}} \qquad (4\text{-}47)$$

and

$$R_{\text{н}} = \rho_T \frac{4 l_{\text{к}}}{\pi d_{\text{к}}^2}: \qquad (4\text{-}48)$$

Since $R_n = U_n/I_n$ and $P_n = I_n U_n$, then solving these two equations simultaneously, we obtain an expression for the cathode diameter:

$$d_{\text{к}} = \sqrt[3]{\frac{4 \rho_T l_{\text{н}}^2}{\pi^2 \beta \eta_{\text{окс}}}}, \qquad (4\text{-}49)$$

where

$$\beta = \frac{d_{\text{окс}}}{d_{\text{к}}}.$$

The ratio of the diameter of an oxide-coated cathode to the diameter of the base may vary from 2 to 3. As a rule, the thickness of the oxide layer is measured not in terms of linear dimensions, but in terms of weight; i.e., by the quantity determining the ratio of the weight of carbonates applied to the base to the rate of the base itself, expressed in per cent. For the most widely utilized method of applying carbonates to bases for filamentary cathodes — the cataphoretic method — the given diameter relationship corresponds to the weights as follows: For a tungsten base, from 28 to 75 per cent, and for a nickel base, from 60 to 160 per cent (the density of the oxide layer is assumed to equal 1.8 g/cm^3).

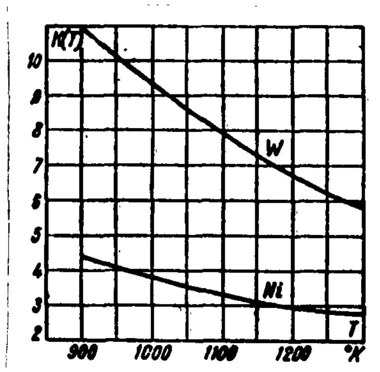

Fig. 4-23. The function K(T) as a function of maximum temperature for tungsten and nickel.

Determining the diameter of the base of the cathode, we may find the length of the ideal cathode corresponding to the given heating parameters:

$$l_{к.ид} = \frac{U_н \pi d_к^2}{I_н \cdot 4\rho_T} .* \qquad (4\text{-}50)$$

In order to determine the actual length of the cathode, it is necessary to introduce a correction allowing for the decrease in cathode resistance owing to the cooled portions. The actual cathode length, corresponding to the distance between the insulators upon which the cathode rests equals:

$$l_к = l_{к.ид} + n\Delta l, \qquad (4\text{-}51)$$

where n is the number of cooled sections of the cathode.

The value for the correction Δl may be found from the relationship

$$\Delta l = K(T)\frac{d_к}{\sqrt{d_{окс}}} . \qquad (4\text{-}52)$$

* [$l_{к.ид}$ - $l_{k.\ id}$ - $l_{katod\ ideal'nyy}$ - $l_{cathode\ ideal.}$]

The function K(T) is shown in the curve of Fig. 4-23 for tungsten and for nickel bases. It is necessary to add still another 5-6 mm to the calculated cathode length on each leg of the cathode; this is necessary for welding to supports and hooks.

For small-diameter cathodes (less than 30 microns) the calculation may be carried out according to the method suggested by S. M. Moshkovich, based upon investigations of the properties of long cathodes, long enough so that the cathodes could be assumed to be ideal.

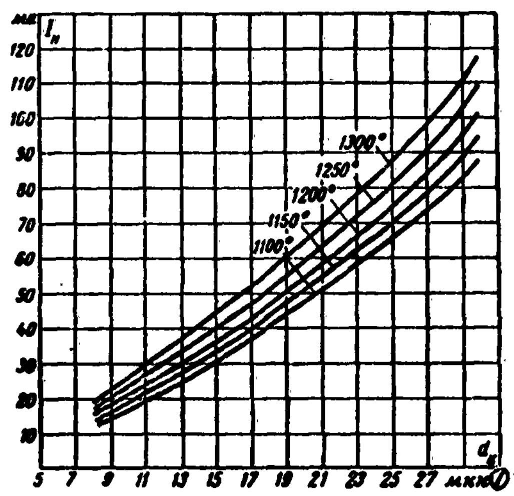

Fig. 4-24. Heating current as a function of cathode-base layer diameter.
1) Microns.

Tungsten wire is used almost exclusively in the bases of thin cathodes. On the basis of experiment, it is assumed that for such cathodes the best relationship between the diameters of the oxide-coated cathode and the base is that for which the thickness of the oxide layer equals the base diameter, i.e., $d_{oks} = 3d_k$.

For such cathodes, experimental investigations have uncovered the dependence of the heating current upon base diameter for various base temperatures, and the dependence of the heating power per centimeter of cathode length upon base diameter for various values of

heating current. Graphs for these relationships are given in Fig. 4-24 and 4-25.

Selecting a maximum base temperature, it is possible, using the curve $I_n = f(d_k)$, to find for this temperature the diameter of cathode base, and then using the curve $P_{nl} = f(d_k)$, in order to get the calculated value of heating current, to find the magnitude of heating current for each centimeter of cathode length.

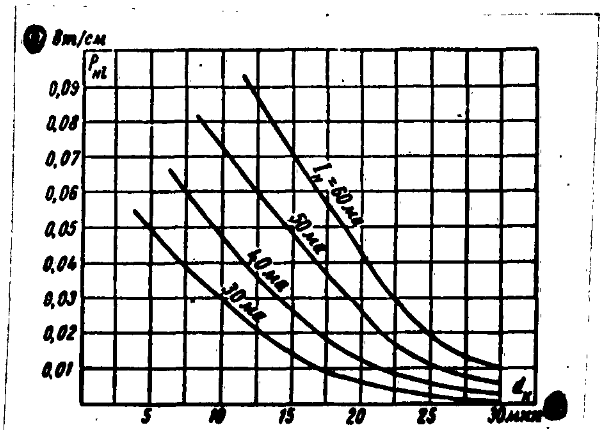

Fig. 4-25. Heating power per unit cathode length as a function of cathode-base-layer diameter. 1) watt/cm; 2) microns.

The total length of an ideal cathode is determined from the relationship

$$l_{k.id} = \frac{P_n}{P_{nl}}. \qquad (4\text{-}53)$$

In determining the true length of a cathode, a correction is introduced for the cooled ends, just as in the case considered previously.

In calculating the diameter of a cathode for a heating current of up to 60 ma, it is possible to utilize Moshkevich's empirical formula

$$d_k[\text{microns}] = 0.2 I_n(\text{ma}) + 6. \qquad (4\text{-}54)$$

4-8. POWER-SUPPLY FOR CATHODE HEATING CIRCUITS.

If the operating conditions for the cathode are properly chosen, and these conditions maintained constant, it is possible to obtain a considerable increase in cathode life, and to provide minimum changes in its other parameters. But the operating conditions of a cathode are determined by its working temperature, which is nearly impossible to monitor. As a rule, the operating regime of a cathode can be monitored only in terms of the heating current, using an ammeter connected into the heating circuit in series with the cathode, or in terms of the heating voltage, using a voltmeter connected in parallel with the cathode. Monitoring of cathode operation in terms of heating current, however, is not equivalent to monitoring with respect to heating voltage, since by the first method, a shorter cathode life is obtained than with the second.

Actually, during the operation of a plain filament, gradual evaporation of the cathode material takes place, resulting in a decrease in diameter. If the heating current is maintained constant during this process, by the end of the useful service life, the cathode temperature increases, since for the same heating current, its density, in view of the decreased cathode diameter, rises. As a result, the rate of evaporation of the cathode material also increases. Thus, the cathode life is noticeably shortened.

If the heating voltage is maintained constant, then as the cathode diameter decreases, its resistance rises, and the heating current drops. As a result, the operating temperature of the cathode will also decrease. Owing to this, the emission current falls somewhat, but cathode life is lengthened. For pure metal cathodes, operated at constant heating voltage, the lifetime normally is greater by a factor of 2-3 than where they are operated with constant heating current.

It is especially important to limit the heating current when the cathode is connected in, since the resistance of the cold cathode is lower than when it is at the operating temperature. Thus, at the instant the connection is made, the starting current proves to be considerably greater than the working current, which may damage the cathode.

Special devices (thermistors, ballast tubes) are used for automatic limiting of the starting current and to maintain the working heating current constant. Starting rheostats are frequently utilized in high-power devices; they are connected into the circuit to limit the starting current, and disconnected after the cathode has heated up.

Either direct current (from a filament battery), or alternating current (from a filament transformer) may be utilized for cathode heating.

When filamentary cathodes are supplied with direct current, an ammeter must be connected into the negative side of the cathode circuit, since in the negative branch of the heating circuit, in addition to the current I_n, there flows in the same direction one-half of the plate current I_a, while at the same time, in the positive branch of the heating circuit, there are the currents I_n and $I_a/2$, which are positive. Thus, the total current flowing in the negative side of the cathode circuit equals the sum of the currents $I_n + I_a/2$, while at the positive end of the cathode we have the difference current $I_n - I_a/2$ (Fig. 4-26). Thus, in order to avoid overheating the negative portion of the cathode, it is necessary to monitor the current in the negative side of the heating circuit.

When a AC supply is used, the voltage drop across the cathode changes periodically both in magnitude and in sign. As a result, the plate voltage and plate current will vary at the same frequency. In

the operation of radio equipment, these changes in plate current create

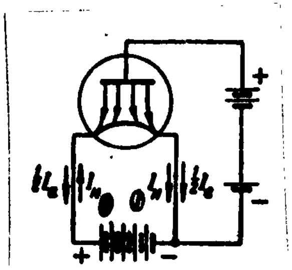

Fig. 4-26. Branching of emission current in cathode heating circuit.
1) I_n.

AC noise, which interferes with normal operation of the equipment. In order to decrease this hum, when filamentary cathodes are supplied with alternating current, the negative terminal of the plate circuit is connected to a center tap on the filament-transformer secondary, or to the center point of a potentiometer connected across the cathode (Fig. 4-27).

With AC supply, the cathode temperature does not remain constant, but changes at a rate of twice the supply frequency. The alternate cooling and heating of the cathode causes changes in the magnitude of the emission current, which also creates AC hum at twice the frequency. This phenomenon is more evident in thin directly heated cathodes, and less evident in massive cathodes and indirectly heated cathodes.

It is also desirable to take into consideration the fact that when heavy oxide-coated indirectly heated cathodes and also glowing oxide cathodes are turned on, it is first necessary to heat up the cathode for a period of 1-3 min, and only then to apply the plate voltage, since the working temperature and emission current reach their normal values for these cathodes 1-3 min later, depending upon the type of cathode. For thin directly heated oxide-coated cathodes, the time required for heating up to the operating temperature all-in-

all amounts to only several seconds.

The ballast tube. The glass envelope of a ballast tube is filled with hydrogen; inside the envelope there is an iron or tungsten wire. Various types of ballast tubes are filled with hydrogen to various pressures (ranging from several tens to 200 mm Ht). Short and thin sections of wire are utilized in the form of straight filaments, and long sections are wound spirally before fastening to the mount.

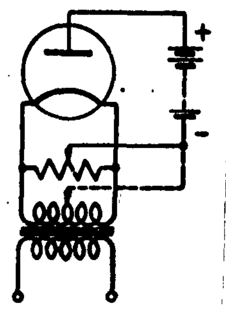

Fig. 4-27. Connection of negative side of plate battery to filament circuit where the filament is supplied with alternating current.

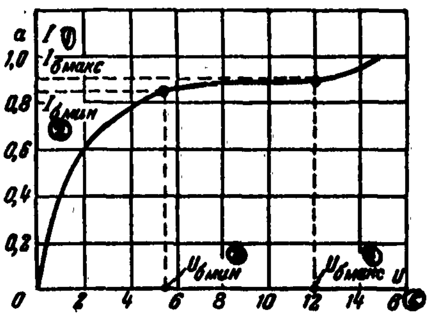

Fig. 4-28. Volt-ampere characteristic of ballast tube. $I_{b\ min}$ and $I_{b\ max}$ – maximum and minimum ballast current; $U_{b\ min}$ and $U_{b\ max}$ – maximum and minimum ballast voltage. 1) $I_{b\ max}$; 2) $I_{b\ min}$; 3) $U_{b\ min}$; 4) $U_{b\ max}$; 5) v.

When current is passed through the ballast tube, the wire is heated up, and its resistance rises. It is possible to select the hydrogen pressure, length and diameter of the wire so that the increase in resistance of the wire when it is heated will be directly proportional to the increase in the voltage applied, and the current flowing in the ballast-tube circuit will remain constant. It is clear that the only type of wire suitable for a ballast tube is a wire having a sufficiently high positive temperature coefficient of resistance. Iron and tungsten are completely satisfactory in this respect; upon heating, there is a considerable rise in their resistance.

The basic parameters of a ballast tube are the magnitude of the current which the tube will maintain constant (stabilization current or ballast current), and the value of the voltage at the beginning and end of stabilization.

If the voltage across a ballast tube is gradually increased, the current flowing through its circuit begins to rise, and then remains practically unchanged, and with a further increase in voltage again rises. Figure 4-28 shows a typical ballast-tube current-voltage relationship. With a change in voltage across the ballast tube from $U_{b\ min}$ to $U_{b\ max}$, the ballast current changes slightly (from $I_{b\ min}$ to $I_{b\ max}$).

Ballast tubes are used for current regulation in filament circuits of receiving and amplifying tubes where there is a possibility of voltage variations in the power source. They may be utilized in either DC or AC circuits, but it should be kept in mind that ballast tubes are thermal devices and, consequently, take time to react, a time amounting to 2-3 min. Thus, a ballast tube will not stabilize current where there are rapid fluctuations in power-supply voltage. When working with alternating current, a ballast tube reacts not to the instantaneous value of the current, but to its effective value.

The basic parameters for several types of ballast tubes are given in Table 4-4.

The lifetime of various types of ballast tubes is different and varies over a very wide range (from 100 to 10,000 hr).

Where a ballast tube is connected in series with an electron tube, it is necessary to make a proper choice of power-supply voltage. For normal operation, the supply voltage is chosen according to the formula

$$U = U_{н} + \frac{U_{н.б} + U_{к.б}}{2},*$$

* [$U_{нб}$ - $U_{n.\ b.}$ - $U_{nachal\ barettirovaniya}$ - $U_{initial\ ballasting}$.]
[$U_{кб}$ - $U_{k.\ b.}$ - $U_{konets\ barettirovaniya}$ - $U_{final\ ballasting}$.]

TABLE 4-4. Ballast-tube parameters.

Обозначение (1)	Ток стабилизации (начальный), ма (2)	Ток стабилизации (в конце срока службы), ма (3)	Напряжение начала и конца стабилизации, в (4)
0,24Б12-18	251—261	250—262	12—18
0,3Б17-35	275—325	270—330	17—35
0,3Б65-135	270—330	265—335	65—135
0,425Б5,5-12	390—460	380—470	5,5—12
0,85Б5,5-12	780—920	760—940	5,5—12
1Б5-9	960—1 040	950—1 050	5—9
1Б10-17	960—1 040	950—1 050	10—17

1) Designation; 2) stabilization current (initial), ma; 3) stabilization current (at end of useful service life) ma; 4) voltage at beginning and end of stabilization, v.

where $U_{n.\ b.}$ and $U_{k.\ b.}$ are the initial and final ballast voltages; U_n is the heating voltage for the tube cathode with which the ballast tube is connected in series.

In this case, the power-supply voltage should not be below the sum $U_n + U_{n.\ b.}$ or above the sum $U_n + U_{k.\ b.}$.

An over-all view of the 0.85Б5.5-12 ballast tube is shown in Fig. 4-29.

The thermistor. Thermistors (thermal resistances) are nondischarge devices which in contrast to ballast tubes have a considerable negative temperature coefficient. Thus, as temperature increases, the resistance of a ballast tube rises, while the resistance of a thermistor drops.

All semiconductors, as is known, have a negative temperature coefficient, but in the manufacture of thermistors, only those semiconductors are suitable that do not have rectifying properties, and that do have excellent resistance constancy with time.

Materials suitable for manufacturing thermistors are cupric oxide, uranium oxide, magnesium-titanium, spinel ($MgTi_2O_4$), frequently called an urdox-resistance, and certain special mixtures of oxides of various

heavy metals (Ni, Fe, Co, Cu, U, etc.), having the negative temperature coefficient and other properties required for thermistors. Certain of the semiconductors, such as, for example, cupric oxide do not react with air, which makes it possible to manufacture cupric oxide thermistors in the form of standard resistances connected directly into the circuit; thermistors manufactured from semiconductors that do react with the air are placed in a vacuum or in an atmosphere of an appropriate gas that does not react with the material of the thermistor.

With an appropriate choice of thermistor size and heat emission from its surface, it is possible to obtain a section of the characteristic over which the voltage is practically invariant with current change, i.e., to obtain a volt-ampere characteristic with a long sloping section, bu. a characteristic completely opposite to that of a ballast tube (Fig. 4-30). As the current flowing through the thermister

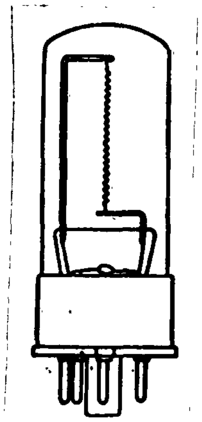

Fig. 4-29. Overall view of 0.85B5.5-12 ballast tube.

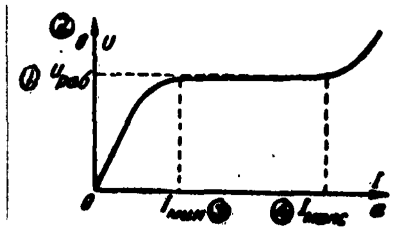

Fig. 4-30. Volt-ampere characteristic of thermister. 1) U_{work}; 2) v; 3) I_{min}; 4) I_{max}.

increases from 0 to I_{min}, the thermistor is slightly heated, owing to which its resistance is nearly unchanged, and the voltage across the thermistor rises as the current increases. When the current increases from I_{min} to I_{max}, the thermistor temperature rises, and its resistance drops. Thus, the voltage across the thermistor remains almost

constant (horizontal working section). With a still further increase in current, the thermistor resistance decreases slightly as the temperature goes up, and the voltage across the thermistor begins to rise.

The working section is the longer the higher the negative temperature coefficient of the thermistor, i.e., the more the thermistor resistance drops as the current increases; here the voltage across the thermistor remains constant. If the change in resistance with varying current is slight, there may not be a well defined working section, which is the case, for example, with cupric oxide thermistors.

The presence of a long working section in the volt-ampere characteristic makes it possible to utilize the thermistor for automatic control of voltage where the current changes from I_{min} to I_{max}. In this case, the thermistor is connected in parallel with the load (Fig. 4-31a). An increase in the power-supply voltage causes an increase in the current through the thermistor, since it heats up and its resistance drops. Thus, the voltage drop across resistor r increases, while the voltage drop across the thermistor T and the load R_n is almost unchanged.

Where a thermistor is connected into a circuit in series with a ballast tube (Fig. 4-31b), almost the entire voltage appears across the thermistor, since its resistance in the cold state is high, and the current in the circuit is small. As current passes, the resistance of the thermistor drops, but the resistance of the ballast tube rises, and the current remains practically constant. Thus, in the given case, the thermistor makes it possible to improve the characteristic of the ballast tube during the instant the circuit is switched on. In addition, the thermistor is also utilized in bridge circuits for measuring the power of high- and super-high-frequency currents. When high-frequency

currents pass through a thermistor connected into one arm of a bridge, it is heated, its resistance drops, and the bridge becomes unbalanced. It is possible to measure the power dissipated in the thermistor by the high-frequency currents by balancing the bridge again. Small thermistors are especially valuable in this connection; they have little inductance or capacitance, and may be connected into super-high-frequency circuits and low powers may be measured accurately with their aid (thousandths of a watt).

Thermistors are utilized for automatic limiting of initial (starting) currents when tube cathodes are connected into a circuit, and for automatic voltage control. In the latter case, the volt-ampere characteristic of the thermistor should have a large horizontal (working) section. Thermistors are also utilized in special circuits for measuring and controling temperature.

The following quantities are included in the parameters of a thermistor:

a) The voltage drop across the thermistor corresponding to the working portion of the volt-ampere characteristic (U_{rab}*);

b) Resistance of the thermistor in the cold state (R_0 at 20°C);

c) The minimum resistance of the thermistor in the heated condition (R_{min}) expressed in per cent of R_0 (at the maximum permissible temperature);

d) The maximum permissible thermistor current, at which the thermistor is heated to the maximum permissible temperature;

e) The minimum thermistor current I_{min};

f) The thermistor time constant determined by the time required to heat up the thermistor.

The values of certain parameters for the most common materials used in the manufacture of thermistors are given in Table 4-5.

* [$U_{раб}$ - U_{rab} - $U_{rabochiy}$ - $U_{working}$.]

TABLE 4-5. Parameters of materials from which thermistors are made.

1 Материалы	R_0, ом 2	3 $R_{мин}/R_0$, %	4 $T_{макс}$, °C	Постоянная времени, сек 5
6 Окись урана . .	5 000—100 000	0,5	600	0,1—10
7 Урдокс	15—100 000	2—4	500	10—600
8 Окись меди . . .	0,5—10 000	10	220	1—100

1) Materials; 2) I_0, ohms; 3) R_{min}/R_0, per cent; 4) T_{max}, °C; 5) time constant, sec; 6) uranium oxide; 7) urdox; 8) cupric oxide.

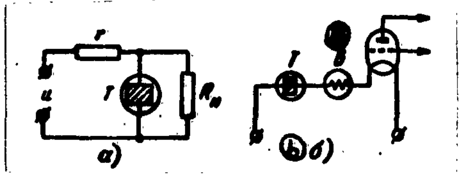

Fig. 4-31. Thermistor circuits. a) Series-connected; b) parallel-connected.

Chapter Five

THE TWO-ELEMENT TUBE — THE DIODE

5-1. THEORY OF SPACE CHARGE. THE THREE-HALVES LAW.

The simplest electronic device is the two-element electron tube, the diode. The diode has an incandescent cathode and a plate which are separated by a vacuum space. Figure 5-1 shows typical diode constructions having directly (a) or indirectly (b) heated cathodes and the symbols used for them.

An electric field forms in the presence of a potential difference between the cathode and the plate. The electrons emitted by the cathode move in the space between plate and the cathode under the influence of this field; if the potential of the plate is higher than the potential of the cathode the electrons move toward the plate, thus creating an electric current in the vacuum. Here the loss of the electrons which travel to the plate is continuously replenished by electrons emitted by the cathode. If the potential of the plate is lower than the potential of the cathode, the electrons emitted by the cathode are returned to the cathode by the electric field and no current flows through the vacuum.

Consequently the diode carries current in only one direction, i.e., is a rectifying device. This property determines the basic purpose of the diode — rectification of alternating currents.

It is important to know for practical purposes how the magnitude of the current passing through the diode depends upon the potential difference between the plate and the cathode. In order to find this dependence, let us consider the electric field between the plate and the cathode. Let us assume for the sake of simplification that the plate and the cathode are parallel infinite surfaces.

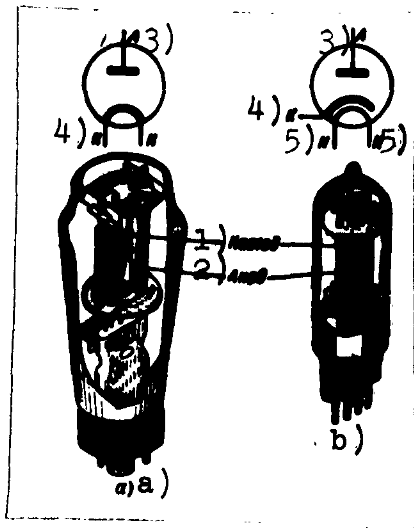

Fig. 5-1. Typical constructions of diodes with direct (a) and indirect (b) heating. 1) Cathode; 2) Plate; 3) Plate; 4) Cathode; 5) Heater.

In the presence of a potential difference between the plate and the cathode and in the absence of any free charges in the space between the plate and the cathode, a uniform electric field whose strength at any point is perpendicular to the surface of the plate and the cathode exists in this space. Surfaces which are parallel to the cathode are equipotential surfaces in this space and consequently potential variation is possible only in the direction perpendicular to the cathode.

Let us plot the potentials of the points under consideration as functions of distance to the cathode. We shall set the potential of the cathode equal to zero.

The field strength at any point of the space is equal to

$$E=\frac{U_a}{r_a} \; * \qquad (5\text{-}1)$$

and consequently the potential of a point situated at a distance **x** from the cathode is equal to

*[U_a = U_a = U_{anod} = U_{plate}.]

61-10

$$U_x = \frac{U_a}{r_a} x. \qquad (5\text{-}2)$$

It is evident from Eq. (5-2) that the dependence of potential on distance is linear and may be graphically represented by a straight line (Fig. 5-2, curve 1).

If the cathode is heated, negative charges (electrons) appear in the space. The presence of a great number of negative charges distributed in the entire space between the cathode and the plate, the so-called space charge, lowers the potential of all points of this space. The more electrons in the space, the more sharply the potential decreases.

The resultant distribution of potentials depends firstly upon the plate potential and secondly on the distribution of the space charge in the space between the plate and the cathode. The greater the density of electrons per unit volume, the greater will be the potential decline in the region of the space in question (Fig. 5-2, curves 2 and 3).

With the plate potential positive, the electrons emitted by the cathode fall into an accelerating electric field and move toward the plate with a velocity which increases as they approach it.

Let us visualize in this space a cylinder with a cross section of 1 cm^2 and an axis perpendicular to the planes of the plate and the cathode (Fig. 5-3). Since the electrons move only in the direction of the X axis, the magnitude of the current through a cross section of this cylinder is constant at any value of x from $x = 0$ to $x = r_a$, where r_a is the distance between the plate and the cathode, and is equal to the product of the charge density ρ_x in this region and the velocity v_x with which the charges move:

$$i = \rho_x v_x. \qquad (5\text{-}3)$$

from which

$$\rho_x = \frac{i}{v_x}, \qquad (5\text{-}4)$$

i.e., the charge density in the space under consideration diminishes as the velocity of their movement increases.

It follows from the above that as the velocity of the electrons increases with their approach to the plate, the density of the space charge diminishes as we recede from the cathode. The greatest charge density is found in the vicinity of the cathode and consequently the potential depression is also greatest close to the cathode.

As a result of the action of the space charge, the potential distribution in the space between the cathode and the plate changes and assumes the form represented by curve 2 in Fig. 5-2. If the

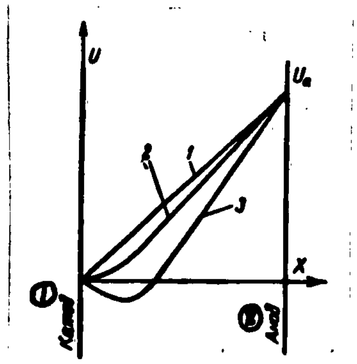

Fig. 5-2. Curves of potential variation in space between plate and cathode with heated and unheated cathodes. 1) Cathode; 2) Plate.

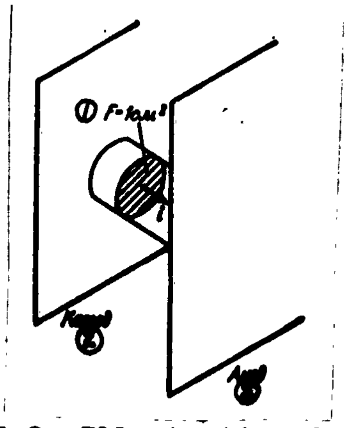

Fig. 5-3. Illustrating derivation of relationship between charge density and velocity. 1) Area = 1 cm^2; 2) Cathode; 3) Plate.

emission of the cathode is sufficiently great, the density of the space charge may be so high that the negative potential created by the space charge in the region close to the cathode is greater than the potential created by the plate voltage. In this case, the zone of the space in which this relationship obtains has a negative poten-

tial. The potential distribution for this case is represented by curve 3, Fig. 5-2.

Let us consider two cases of potential distribution in the space between the plate and the cathode in the presence of a distributed charge. The former case corresponds to curve 2 and is distinguished by a continuous potential increase from the cathode toward the plate. Here any electron coming from the cathode falls into an accelerating field and will necessarily reach the plate and, consequently, the plate current is equal to the cathode emission current. This mode of current passage is known as the saturation mode.

In the second case of potential distribution, which corresponds to curve 3, a potential minimum is formed between the plate and the cathode and divides the whole space into two segments, a segment between the cathode and the potential minimum and a segment between the potential minimum and the plate. Electrons emitted by the cathode fall into a retarding electric field acting on the segment between the cathode and the potential minimum. Only those electrons which, having overcome this retarding field, pass beyond the potential minimum can reach the plate. Electrons whose initial energy is not sufficient to overcome the potential minimum return to the cathode. The plate current is smaller than the cathode emission current in this case, and depends on the plate potential at a given emission magnitude. This mode is called the space-charge mode and is typical of the majority of electronic devices.

At constant cathode emission, the potential distribution varies with the plate voltage. Figure 5-4 shows curves of potential distribution for various plate voltages. When the plate voltage is zero, the whole space between the plate and the cathode has a negative potential. An increase in plate voltage shifts the potential

minimum toward the cathode, and this is accompanied by a decline in the absolute magnitude of the potential minimum. Finally, at a sufficiently high plate voltage the potential of all points of the space becomes positive with the onset of the saturation mode.

In the space-charge mode, the electrons going to the plate from the space charge are replenished by electrons entering the space charge from the cathode. In order to determine the relationship between the magnitude of the plate current and plate voltage, let us consider the region of the plate-to-cathode space between the potential minimum and the plate. In order to simplify the derivation, let us stipulate the following conditions:

1. The emission of the cathode is unlimited; this provides replenishment of the space charge by electrons for any selected current to the plate.

2. Electrons leave the cathode with an initial velocity of zero.

3. The potential minimum is situated at the cathode. It follows from this condition that: a) the distance between the plate and the potential minimum is made the same as the distance between the plate and cathode: b) the potential difference between the plate and the potential minimum is set equal to the plate voltage; c) the field strength at the cathode is considered to be zero.

To derive the magnitude of the plate current as a function of plate voltage, let us use the Ostrogradsky-Gauss theorem, according to which the electric flux through a closed surface is proportional to the magnitude of the charge inside this closed surface:

$$\Sigma E_n \Delta S = 4\pi q. \qquad (5\text{-}5)$$

where q is the magnitude of the charge, and E_n is the field-strength component normal to a surface element ΔS.

Let us take a volume in the form of a cylinder whose axis is perpendicular to the planes of the plate and the cathode and whose bases consist of parts of the plate and the cathode themselves (Fig. 5-5). Let us denote the area of the cross section by F_a. Since the field strength is directed along the axis of the cylinder and its lines of force consequently do not intersect the lateral surface, while, in accordance with the conditions adopted, the field strength at the cathode is zero, the entire flux is determined for any cross section of the cylinder at a distance x from the cathode by the product of the field strength in that cross section by the area F_a. Then Eq. (5-5) takes the form

$$E_x F_a = 4\pi q, \tag{5-6}$$

where q is the magnitude of the charge situated in the volume of the cylinder between the cathode and the selected cross section.

Let us denote by t the time in which all of this charge passes through the cross section of the cylinder. Then the current

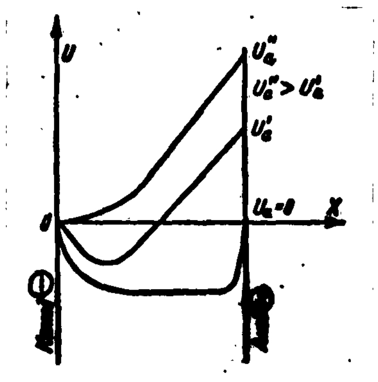

Fig. 5-4. Potential distribution as a function of plate voltage. 1) Cathode; 2) Plate.

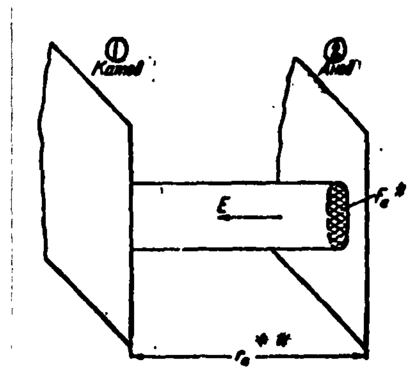

Fig. 5-5. Illustrating derivation of equation of the three-halves law. 1) Cathode; 2) Plate.

*[F_a = F_a = F_{anod} = F_{plate}.]
**[r_a = r_a = r_{anod} = r_{plate}.]

through the cylinder is

$$I=\frac{q}{t}. \tag{5-7}$$

According to Newton's law, $F=m\frac{dv}{dt}$, and

$$m\frac{dv}{dt}=eE, \tag{5-8}$$

where dv/dt is the rate of velocity change with time, i.e., the acceleration of the electron, and the product eE determines the force acting upon the electron in the electric field.

We substitute the value of E from (5-6) in Eq. (5-8):

$$m\frac{dv}{dt}=\frac{4\pi eq}{F_a} \tag{5-9}$$

and, substituting the value of q from (5-7), we obtain

$$m\frac{dv}{dt}=\frac{4\pi eI}{F_a}t, \tag{5-10}$$

from which it follows that

$$m\,dv=at\,dt, \tag{5-11}$$

where $a=\frac{4\pi eI}{F_a}$.

Integrating this equation, we obtain

$$mv=a\frac{t^2}{2}+C_1; \tag{5-12}$$

the quantity $a=\frac{4\pi eI}{F_a}$ is constant, since the current remains unchanged in any cross section of the cylinder under consideration. Since at t = 0 (the initial moment of time) $v_0 = 0$ by condition, then $C_1 = 0$.

Velocity may be expressed as the first derivative of distance with respect to time, i.e., v = dx/dt. Thus Eq. (5-12) may be written in the following form:

$$m\frac{dx}{dt}=a\frac{t^2}{2} \tag{5-12a}$$

or

$$m\,dx=a\frac{t^2}{2}dt. \tag{5-12b}$$

Integrating Eq. (5-12b), we obtain

$$mx = a\frac{t^3}{6} + C_2. \qquad (5\text{-}13)$$

Since $x = 0$ at $t = 0$ (the initial moment when electrons move from the cathode) the integration constant C_2 is also zero. Consequently,

$$mx = a\frac{t^3}{6}. \qquad (5\text{-}13a)$$

This equation holds true for any value of x. Substituting $x = r_a$ in it, we get an expression determining the transit time of the electrons from the cathode to the plate:

$$\tau = \left(\frac{6mr_a}{a}\right)^{1/3}. \qquad (5\text{-}14)$$

The magnitude of the transit time may also be gotten by dividing Eq. (5-13a) by Eq. (5-12), remembering that $C_1 = 0$

$$t = \frac{3x}{v}. \qquad (5\text{-}15)$$

At the point $x = r_a$, i.e., when the electrons arrive at the plate, their velocity is determined by the plate potential. The relation $v = \sqrt{2eU/m}$ where U is the potential difference crossed by the electrons, was derived earlier. Substituting the quantity r_a for x and the value of the velocity of the electrons passing to the plate in Eq. (5-15), we obtain

$$\tau = \frac{3r_a}{\sqrt{2\frac{e}{m}U_a}}. \qquad (5\text{-}16)$$

This is a second expression for the transit time of the electrons. Equating the right members of Eqs. (5-16) and (5-14) and replacing the quantity a by its value $(4\pi eI/F_a)$ we obtain

$$\left(\frac{6mr_aF_a}{4\pi eI}\right)^{\frac{1}{3}} = \frac{3r_a}{\left(2\frac{e}{m}U_a\right)^{\frac{1}{2}}}. \qquad (5\text{-}17)$$

It is possible to determine from this equation the relationship between the magnitude of the plate current and the plate voltage for a planar diode:

$$I = \frac{1}{9\pi}\sqrt{2\frac{e}{m}}\frac{F_a}{r_a^2}U_a^{3/2}. \qquad (5\text{-}18)$$

Equation (5-18) shows that a diode plate current that is limited by the space charge is directly proportional to the potential difference between the plate and the cathode raised to the three-halves power.

Substituting the values of π and e/m, we obtain

$$I_a = 2{,}33\cdot 10^{-6}\frac{F_a}{r_a^2}U_a^{3/2}\ [a],^{*}\quad [\text{amp}] \qquad (5\text{-}19)$$

where U_a is expressed in volts.

The analogous calculation for a cylindrical system, in which the plate and the cathode are coaxial cylindrical surfaces (Fig. 5-6), leads to the following expression for the plate current:

$$I_a = 2{,}33\cdot 10^{-6}\frac{2\pi r_a l}{r_a^2\beta^2}U_a^{3/2}. \qquad (5\text{-}20)$$

where r_a is the radius of the plate; $\underline{l}$ is the effective length of the plate; β^2 is a function of the ratio of the plate radius to the cathode radius.

The product $2\pi r_a l$ is the size of the plate surface absorbing current. In the general case, it is possible to express the relationship between plate current and plate voltage by the following equation:

$$I_a = GU_a^{\frac{3}{2}}. \qquad (5\text{-}21)$$

This equation is called the three-halves law (the Child-Langmuir law).

The coefficient $G = 2{,}33\cdot 10^{-6}\frac{F_a}{r_a^2\beta^2}$ depends on the shape and dimensions of the electrodes of the tube. In the case of the plane system, the function β^2 is unity. In the cylindrical system of electrodes, the quantity β^2 as a function of the ratio r_a/r_k is presented by Table 5-1 and represented graphically in Fig. 5-7. For $r_a/r_k \geq 10$, we may consider $\beta^2 = 1$.

*[I_a = I_a = I_{anod} = I_{plate}.]

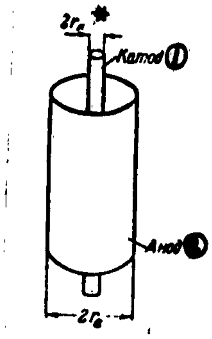

Fig. 5-6. Cylindrical arrangement of diode electrodes. 1) Cathode; 2) Plate.

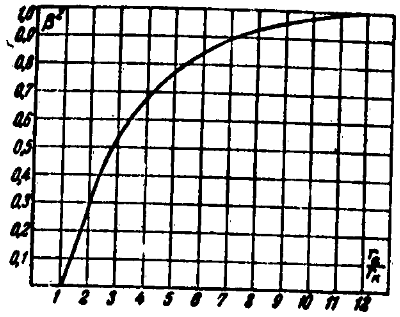

Fig. 5-7. Curve of function β^2.

TABLE 5-1

$x = \frac{r_a}{r_k}$	$\beta^2(x)$	$x\beta^2(x)$	$x = \frac{r_a}{r_k}$	$\beta^2(x)$	$x\beta^2(x)$
1,00	0	0	3,2	0,5526	1,7683
1,10	0,00842	0,00926	3,4	0,5851	1,9899
1,15	0,0175	0,0201	3,6	0,6148	2,2133
1,2	0,02875	0,0345	3,8	0,6420	2,4396
1,3	0,0559	0,0727	4,0	0,6671	2,6684
1,4	0,0867	0,1214	4,2	0,6902	2,8988
1,5	0,1193	0,1790	4,4	0,7115	3,1306
1,6	0,1525	0,2440	4,6	0,7313	3,3640
1,7	0,1854	0,3152	4,8	0,7496	3,5981
1,8	0,2177	0,3919	5,0	0,7666	3,8330
1,9	0,2491	0,4733	5,2	0,7825	4,0690
2,0	0,2793	0,5586	5,4	0,7973	4,3054
2,1	0,3083	0,6474	5,6	0,8111	4,5422
2,2	0,3361	0,7394	5,8	0,8241	4,7798
2,3	0,3626	0,8340	6,0	0,8362	5,0172
2,4	0,3879	0,9310	6,5	0,8635	5,6128
2,5	0,4121	1,030	7,0	0,8870	6,2090
2,6	0,4351	1,131	7,5	0,9074	6,8055
2,7	0,4571	1,234	8,0	0,9253	7,4024
2,8	0,4780	1,338	8,5	0,9410	7,9985
2,9	0,4980	1,444	9,0	0,9548	8,5932
3,0	0,5170	1,551	9,5	0,9672	9,1884
			10,0	0,9782	9,7820

The size of the theoretical surface of the plate or, as it is usually called, the effective surface of real diodes whose plates are longer than the cathode, is computed on the basis of the effective length of the cathode for all electrode systems.

In the case of a plane-system diode with a directly heated cathode in the form of several loops (Fig. 5-8), the magnitude of the effective plate surface may be determined with sufficient accuracy by the method proposed by Kuzunoz. This method consists in taking

*[$2r_к = 2r_k = 2r_{katod} = 2r_{cathode}$.]

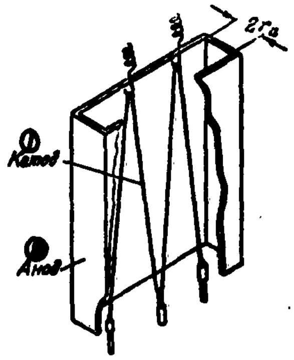

Fig. 5-8. Diode system: flat plate and looped cathode. 1) Cathode; 2) Plate.

the projection of the cathode onto the plate as the effective electron-absorbing plate surface, with the width of the cathode made equal to twice the distance between the cathode and the plate. When the plate is two-sided, the size of this projection must be doubled.

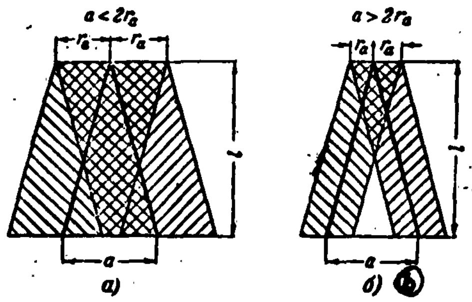

Fig. 5-9. Illustrating determination of effective plate surface in diode having looped cathode.

If the distance between the plate and the cathode is greater than half the distance between the points of attachment of the cathode loops, the size of the unknown surface may be regarded as the area of the trapezoid (Fig. 5-9a). When $r_a < a$ (Fig. 5-9b) the areas of the triangles formed between the crosshatched regions must be subtracted from the area of the trapezoid.

The remaining diode forms can be reduced to the three systems

analyzed for calculation: the plane system, the cylindrical system, and the plane system with looped cathode.

5-2. VOLT-AMPERE CHARACTERISTIC OF DIODE.

The circuit presented in Fig. 5-10 is used for the practical determination of the current flowing through the diode — the plate current — as a function of plate voltage. The plate circuit of the tube is fed by the direct-current source E_a* through the potentiometer P. Heating the cathode is effected by the source E_n** through the regulating rheostat. The magnitude of the plate current is measured by a milliammeter connected into the plate circuit and the magnitude of the plate voltage by the voltmeter V_a***, which is connected in parallel with the potentiometer. Connecting a plate voltmeter directly between the plate and the cathode of the tube is inadmissible, since then the milliammeter will show the sum of

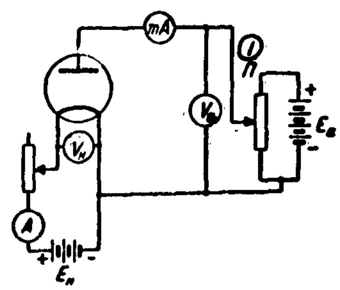

Fig. 5-10. Diagram for study of diode. 1) Potentiometer.

the plate current of the tube and the current drawn by the voltmeter itself. Usually these currents are comparable in magnitude and the error of the milliammeter readings when a plate voltmeter is improperly connected may be quite large.

*[E_a = E_a = E_{anod} = E_{plate}.]
**[E_H = E_n = E_{nakala} = E_{heater}.]
***[V_a = V_a = V_{anod} = V_{plate}.]

Heater voltage is measured by the voltmeter V_n and the heater current by the ammeter. Using this system, it is practically possible to record the volt-ampere characteristic of the diode, i.e., a curve representing the plate current of the tube as a function of plate voltage. Theoretically, the volt-ampere characteristic of the diode must, according to the three-halves law, be a semicubic parabola. Figure 5-11 presents the theoretical characteristic of a two-element tube of type 6Kh2P (curve I) and its real characteristic (curve II).

As is clear from the curves, the real volt-ampere characteristic differs from the theoretical. The cause of this is that in deriving the three-halves law, we made a series of simplifying assumptions which are not met in the real diode.

Let us consider the basic factors bringing about the deviation of the real volt-ampere curves of the diode from the theoretical.

1. <u>Influence of magnitude of cathode emission</u>. In the derivation of the three-halves law, it was assumed that the cathode

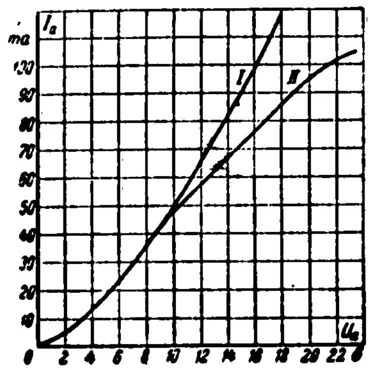

Fig. 5-11. Theoretical (I) and real (II) characteristics of 6Kh2P [REGE] diode.

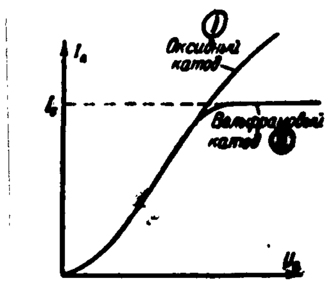

Fig. 5-12. Typical volt-ampere characteristics of diodes with regions of saturation.
1) Oxide-coated cathode; 2) Tungsten cathode.

gives unlimited emission. In practice, the increase in the magnitude of the plate current is limited by the emissive power of the cathode. When the plate current reaches a magnitude equal to the cathode emission current, it undergoes practically no further increase; saturation intervenes. Saturation corresponds to the nearly horizontal region of the curve, as represented in Fig. 5-2. However, distinct saturation is observed only in tungsten or carburized cathodes. In tubes with oxide-coated cathodes, no clearly defined region of saturation can be observed, and the increase in current with increasing plate voltage continues, although more slowly than on the working segment of the characteristic. This phenomenon is related to the so-called Schottky effect, which manifests itself in increased emission from activated surfaces with increasing electric-field strength at the surface.

In real tubes, the transition from the ascending branch of the curve to the region of saturation occurs gradually, as shown in Fig. 5-13; this is explained by the temperature nonuniformity of cathodes in real tubes. Since the emission from cooled regions of the cathode is smaller than from the hotter regions, the saturation intervenes in cooled regions at lower plate-voltage values. For this reason, the current rise begins to slow down at plate-voltage values smaller than the voltage corresponding to complete saturation.

2. Influence of nonequipotentiality of cathode. In deriving the three-halves law, the cathode potential was considered to be uniform and equal to zero over the whole surface. This condition holds true for tubes with indirectly heated cathodes; it is not met in the case of a directly heated cathode. The heater current flowing in the cathode creates a voltage drop across it equal to the heater voltage. As a result, the potentials of different points on the

cathode are found to be unequal and, consequently, the potential difference between the plate and the various regions of the cathode differs from the plate voltage. Figure 5-14 gives a schematic representation of a tube with direct heating in which the negative pole of the heater-current source is connected to the common point 0. In this circuit, the source of the plate voltage U_a and the source of the heater voltage U_n are counterconnected in series. The potential difference between the plate and point 0 of the cathode is equal to plate voltage U_a and that between the plate and point P (the other end of the cathode) is equal to the difference between the plate voltage and the heater voltage. In the case under consideration, therefore, the potential difference between the plate and a given point on the cathode (except for their common point) is less than the plate voltage and the magnitude of the plate current is less than the theoretical. For this reason, the volt-ampere characteristic lies below the theoretical. When the positive pole of the heater-voltage source is connected to the common point, the potentials of all points of the cathode are lower than that of the common point. The potential distribution over the cathode for this case is represented by the broken curve in Fig. 5-14. When the positive pole of

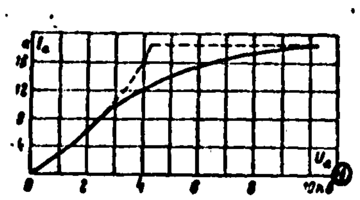

Fig. 5-13. Influence of cathode-temperature nonuniformity on the transition of volt-ampere characteristic of diode to region of saturation (V1-0.1/30). 1) kw.

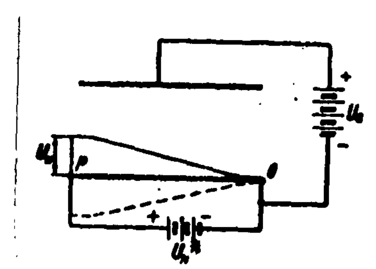

Fig. 5-14. Potential distribution over directly heated cathode.

*[U_H = U_n = U_{nakala} = U_{heat}.]

the heater-voltage source is connected to the common point, the resultant potential difference between the plate and the cathode is greater than the plate voltage, and the volt-ampere curve must run above the theoretical characteristic.

In ordinary diode circuits, the negative pole of the heater-voltage source is connected to the common point, and therefore in designing these tubes we should proceed from the relationships corresponding to this case.

The calculation of the plate current of the diode with a directly heated cathode may be made by the following formula with a correction for the nonequipotentiality of the cathode:

$$I = \frac{2}{5} \cdot 2{,}33 \cdot 10^{-6} \frac{F_{\text{a}}}{r_{\text{a}}^{2} \beta^{2}} U_{\text{н}}^{\frac{3}{2}} f\left(\frac{U_{\text{a}}}{U_{\text{н}}}\right). \qquad (5\text{-}22)$$

Values of the function $f(U_a/U_n)$ are presented in Table 5-2. When the plate voltage exceeds the heater voltage by more than 40 times, the influence of the cathode-potential nonuniformity may be disregarded.

TABLE 5-2.

$\frac{U_{\text{a}}}{U_{\text{н}}}$	$f\left(\frac{U_{\text{a}}}{U_{\text{н}}}\right)$	$\frac{U_{\text{a}}}{U_{\text{н}}}$	$f\left(\frac{U_{\text{a}}}{U_{\text{н}}}\right)$	$\frac{U_{\text{a}}}{U_{\text{н}}}$	$f\left(\frac{U_{\text{a}}}{U_{\text{н}}}\right)$	$\frac{U_{\text{a}}}{U_{\text{н}}}$	$f\left(\frac{U_{\text{a}}}{U_{\text{н}}}\right)$
0,25	0,031	1,50	2,570	4,0	16,5	10,0	74,0
0,50	0,177	2,00	4,650	5,0	24,1	15,0	138,0
0,75	0,414	2,50	7,130	6,0	32,5	20,0	211,0
1,00	1,000	3,00	9,940	8,0	53,0	40,0	622,0

If the heater circuit of the tube is fed by alternating current, the potential difference between the plate and the cathode will be alternately higher and lower than the plate voltage. These oscillations occur with the frequency of the alternating current. As a result, the plate current is not a steady current but a pulsating current. This pulsating current may be represented as the

sum of a constant component I_{a0} and an alternating component having a frequency equal to that of the filament current. The magnitude of the pulsations, which is given by the ratio of the amplitude of the alternating component to the constant component of the current expressed in per cent, is larger for higher filament voltages and lower plate voltages.

In order to suppress pulsations in circuits with alternating-current heater supply to the diode, we use the connection shown in Fig. 4-27: the centerpoint of the transformer filament winding, which has a potential approximately equal to the potential in the center of the cathode, serves as the common point.

Then the potential of one of the ends of the cathode is higher at any moment of time than the potential of the common point by half the instantaneous value of the filament voltage, and the potential of the other end is smaller by the same amount. At a given instant of time, the increase in the potential difference between the plate and one half of the cathode is compensated by a decrease in the potential difference between the plate and the second half of the cathode. This hookup reduces the oscillation of the plate current markedly, but a small amount of pulsation still remains. This pulsation is due to the inevitable asymmetry of the emission properties and dimensions of the cathode halves and also to the non-linearity of the characteristic. Only in tubes with indirectly heated cathodes, whose potential is uniform over the entire surface, is it possible to eliminate completely the pulsations stemming from the oscillation of the cathode potential.

3. Influence of emission properties of cathode on plate current of diode. There is no connection with the emission properties of the cathode in any of the formulas considered for the calculation of the diode plate currents. Tube-manufacturing practice has shown,

however, that the parameters of the diode, including plate current, depend essentially on the magnitude of the cathodic emission and, consequently, on the technology of tube construction. This dependence stems from the following causes.

First, in deriving the three-halves law, it was assumed that the potential minimum was situated at the cathode and the potential difference between the plate and the space charge was equal to the plate voltage. In reality, it is necessary to take into account the distance from the cathode to the potential minimum and the magnitude of the potential minimum, i.e., quantities depending on the emission of the cathode. Taking this correction into account, the formula for the plate current assumes the following form:

$$I_a = 2{,}33 \cdot 10^{-6} \frac{F_a}{(r_a - x_m)^2 \beta^2} (U_a - U_m)^{\frac{3}{2}}, \tag{5-23}$$

where x_m is the distance from the cathode to the potential minimum and U_m is the magnitude of the potential minimum. In practice, the distance from the cathode to the potential minimum at ordinary plate voltages is a few microns and the magnitude of the potential minimum is of the order of hundredths or tenths of a volt. For this reason, these quantities should be taken into consideration only when the distance between the cathode and the plate is short and the plate voltage is small.

Secondly, the cathodes of real tubes have different emission properties in different regions. Poorly activated regions of the cathode possess little emission and work in the saturation mode, and this lowers the magnitude of the plate current as compared with the theoretical.

It follows from the above that the magnitude of the diode plate current depends to a high degree on the fabrication process. This is also true for other types of electron tubes.

5-3. PARAMETERS OF THE DIODE.

For practical utilization of the diode, it is important to know how the plate current varies with varying plate voltages, i.e., how steeply the volt-ampere characteristic rises. Since the shape of the volt-ampere characteristic approximates a semicubic parabola, the steepness of its ascent differs for different plate voltages. For this reason, it is possible in considering any region of the curve to speak of an average slope of the curve in that region and replace it by a straight line. Let us consider the volt-ampere curve of a diode (Fig. 5-15). At a given value of the plate voltage U_a', the plate current has a value of I_a'. A change in the plate voltage

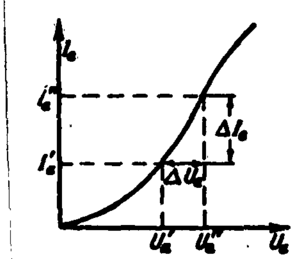

Fig. 5-15. Illustrating definition of the slope of a diode.

to U_a'' brings about a change in the plate current to I_a''. The diode's average slope in the segment from $U_a = U_a'$ to $U_a = U_a''$ is defined as the ratio of the plate-current increment to the plate-voltage increment that produced it:

$$S_{cp} = \frac{I_a'' - I_a'}{U_a'' - U_a'} = \frac{\Delta I_a}{\Delta U_a} \text{ [ma/v]*} \qquad (5\text{-}24)$$

Consequently, the average slope of the diode shows by how much the diode plate current changes when the plate voltage changes by 1 v in a given plate-voltage interval. As this interval is made smaller, the value of average slope approaches the true value. At the

*[$S_{cp} = S_{sr} = S_{srednyy} = S_{average}$.]

limit, with ΔU_a tending to zero, we obtain the value of the slope at the point in question:

$$S = \lim_{\Delta U_a \to 0} \frac{\Delta I_a}{\Delta U_a} = \frac{dI_a}{dU_a}. \tag{5-25}$$

Expressing the value of the plate current according to the three-halves law, $I_a = GU_a^{3/2}$, in Eq. (5-25) and differentiating, we get an expression for the slope at the given point:

$$S = \frac{d(GU_a^{3/2})}{dU_a} = \frac{3}{2} GU_a^{\frac{1}{2}}. \tag{5-26}$$

Replacing the quantity G by its value $2{,}33 \cdot 10^{-6} \frac{F_{a.\text{эф}}}{r_a^2 \beta^2}$ *, we finally obtain

$$S = 3{,}5 \cdot 10^{-6} \frac{F_{a.\text{эф}}}{r_a^2 \beta^2} U_a^{\frac{1}{2}} \text{ [amp/v]} \tag{5-27}$$

Comparing Eq. (5-26) with the formula of the three-halves law, we may obtain the relationship between the magnitude of the plate current and the slope at any point of the curve:

$$\frac{I_a}{S} = \frac{GU_a^{3/2}}{\frac{3}{2} GU_a^{\frac{1}{2}}}, \tag{5-28}$$

from which

$$I_a = \frac{2}{3} SU_a. \tag{5-29}$$

Since the characteristic of the real diode differs from the theoretical, the diode slope may be calculated for practical purposes directly from the characteristic registered from a real tube. The slope at a given point (at a given plate voltage) may be determined by drawing the tangent to that point of the curve. Such a construction is shown in Fig. 5-16. The slope is determined in this case by the slope α of the tangent calculated in the scale on which the curve is constructed.

A diode connected into an alternating-current circuit may be

*[$F_{a.\text{эф}}$ = $F_{a.ef}$ = $F_{\text{anode, effectivnaya}}$ = $F_{\text{plate, effective}}$.]

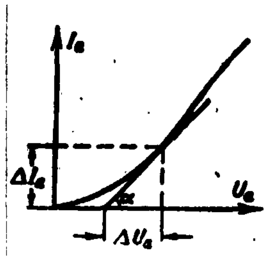

Fig. 5-16. Determination of the slope at a given point of curve with the aid of a tangent.

regarded as a resistance across which a voltage drop forms. Since the relation between the diode current and the voltage drop across the diode is not linear, a plate-voltage change corresponds to different changes in the voltage drop across the tube in different regions of the diode characteristic. The relationship in which the voltage drop varies with plate-current changes is characterized by a second parameter of the diode, the internal resistance to alternating current. If the plate-current change ΔI_a corresponds to a plate-voltage change ΔU_a, the internal resistance of the tube is

$$R_i = \frac{\Delta U_a}{\Delta I_a}. \tag{5-30}$$

On comparing Eq. (5-30) and (5-24) we find that the internal resistance of a diode is the reciprocal of the slope

$$R_i = \frac{1}{S}. \tag{5-31}$$

For this reason, the slope is sometimes measured in conductivity units: mhos or micromhos

The potential difference between the plate and the cathode of a diode forms an electric field which expends energy in carrying the electrons from the cathode to the plate. The energy expended in the transfer of one electron is equal to the kinetic energy acquired by the electron on its path to the plate:

$$w_{\kappa}=\frac{mv^2}{2}, ^{*} \qquad (5\text{-}32)$$

where v is the velocity of the electron arriving at the plate. Assuming the initial electron velocities to be zero, we obtain

$$\frac{mv^2}{2}=eU_a. \qquad (5\text{-}33)$$

As the fast-moving electrons hit the plate, all their kinetic energy is given up to the plate and converted into energy of thermal agitation of the atoms and molecules of the plate's material. The energy coming to the plate in 1 sec, i.e., the power developed at the plate, is equal to the product of the energy of one electron by the number of electrons hitting the plate in the course of 1 sec:

$$P_a=Nw_{\kappa}=NeU_a, \qquad (5\text{-}34)$$

where P_a is the power developed at the plate and N is the number of electrons coming to the plate in 1 sec. The product Ne is the quantity of charge coming to the plate in the course of 1 sec, i.e., the plate current of the diode. Consequently,

$$P_a=I_aU_a. \qquad (5\text{-}35)$$

Thus, all of the electric power lost in the diode is liberated at the plate in the form of heat.

In the steady-state mode which corresponds to a constant plate temperature, the entire power coming to the plate must be given off by the plate to the surrounding space. Plates situated inside bulbs dissipate power by emission and by conduction of heat through the components supporting the plate. The major part of the power is emitted, and the power withdrawn by the supporting components is often disregarded in calculation.

According to the Stefan-Bolzmann law, the power radiated by a heated surface is

*[w_{κ} = w_k = $w_{\text{kineticheskaya}}$ = w_{kinetic}.]

$$P_{изл} = \varepsilon\sigma T^4 F_{изл},^{**} \qquad (5\text{-}36)$$

where F_{izl} is the area of the emitting surface.

It is evident from Eq. (5-36) that the temperature of the plate is proportional to $\sqrt[4]{P_a}$, considering that $P_a = P_{izl}$.

The admissible plate working temperature is limited firstly by the possibility of its becoming distorted at high temperature and secondly by the phenomenon of additional heating of the cathode by the emission from the internal surface of the plate. This phenomenon is particularly dangerous in tubes with oxide-coated cathodes, since overheating of an oxide-coated cathode brings about its rapid disintegration.

Two methods may be used in order to increase the power dissipated by the plate without increasing its temperature: blackening the surface, which raises the emission coefficient, and by consruction of plates with increased emission surfaces achieved by additional ribbing.

It should be noted that the working temperature of the tube plate must not exceed the plate heating temperature at the time of tube evacuation. Otherwise additional evolution of gas from the plate may occur and lead to deterioration of the vacuum in the tube and tube failure.

In case the necessary power cannot be dissipated by emission alone, the plates are air- or water-cooled.

5-4. APPLICATIONS OF DIODES.

a) Rectification of low-frequency alternating currents. The basic application of diodes is the rectification of alternating currents (such diodes are called kenotrons). The circuit of the simplest

*[$P_{изл} = P_{izl} = P_{izluchenia} = P_{emission}$.]
**[$F_{изл} = F_{izl} = F_{izluchenia} = F_{emission}$.]

diode rectifier is shown in Fig. 5-17. In this circuit, the diode is connected in series with the load R in the secondary-winding circuit of the transformer. In the half-cycle when the plate is positive with respect to the cathode, a current flows through the tube and the load, and the transformer voltage is distributed between the tube and the load. In the positive half-cycle, therefore, the plate voltage is lower than the voltage of the secondary transformer winding. No current flows through the diode in the negative half-cycle and the voltage drop across the load is zero. For this reason, the whole transformer voltage falls on the tube in the negative half-cycle.

Figure 5-18 shows the curves of variation of the transformer voltage, the plate current and the plate voltage for a tube working in the circuit under consideration. As is evident from these curves, a pulsating direct current flows through the load. In order to level the pulsations, a rather large capacitance is connected in

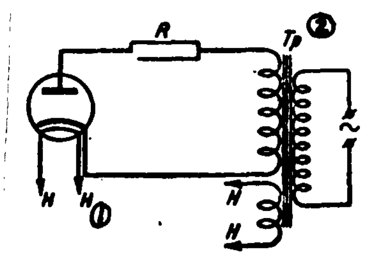

Fig. 5-17. Simplest circuit of half-wave rectifier. 1) Heater; 2) Transformer.

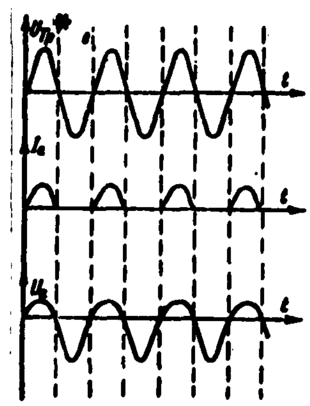

Fig. 5-18. Curves of voltage and current in half-wave rectifier without a filter.

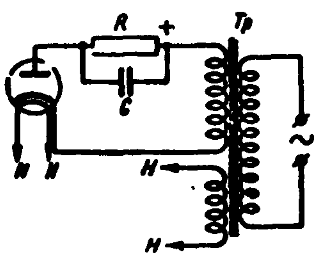

Fig. 5-19. Half-wave rectifier with capacitive output.

*[$U_{Tp} = U_{Tr} = U_{Transformator} = U_{Transformer}$.]

parallel with the load (Fig. 5-19) to serve as a smoothing capacitance. In the positive half-cycle, the capacitor is charged to a potential equal to the maximal voltage drop across the load. When the transformer voltage decreases, the capacitor begins to discharge across the load, supporting a current in it. If the capacitance of the capacitor is large, the discharge time of the capacitor across the load may be much larger than the period of the voltage to be rectified and by the time the voltage across the transformer increases in the subsequent period, the voltage on the capacitor remains quite high. For this reason, the pulsation of the voltage in the load becomes relatively small, as shown by the curve in Fig. 5-20.

This diagram shows the time curves of the voltage across the secondary winding of the transformer, the tube plate current, the plate voltage, and the voltage across the load, which is equal to the voltage across the parallel-connected capacitor.

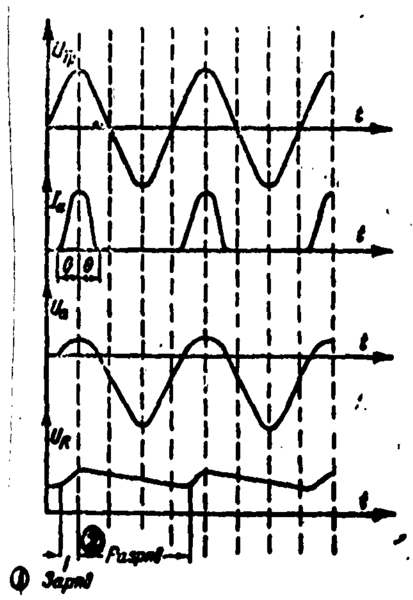

Fig. 5-20. Curves of the working process in half-wave rectifier with capacitive output. 1) Charge; 2) Discharge.

At any moment of time, the voltage across the tube is equal to the algebraic sum of the transformer secondary voltage and the

capacitor voltage.

In the positive half-cycle, the transformer and capacitor are opposed, and tube current flows only when the transformer voltage exceeds the capacitor voltage; the diode current then rapidly rises to the maximum value, the capacitor begins to discharge, and the current begins to decline again in the second quarter of the cycle. When the plate current begins to decline, the capacitor begins to discharge slowly at first, since part of the load current flows through the tube, and then when the voltage across the transformer becomes equal to zero, the discharge proceeds exponentially. Thus, the plate current flows during only part of the positive half-cycle, which is given by the angle 2θ. The pulsations are the smaller the greater the output capacitance of the rectifier and the greater the resistance of the load, i.e., the smaller the load current.

The half-wave rectifier circuit is used comparatively rarely, since it enables us to use only half the transformer power and the pulsations of the rectifying voltage that arise in half-wave rectification are hard to level, even by using massive filters.

A considerably more common circuit is the full-wave rectifier, one variant of which is represented in Fig. 5-21.

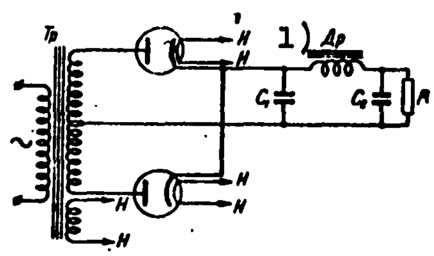

Fig. 5-21. Full-wave rectifier circuit with pi-section filter. 1) Choke.

The secondary winding of the transformer has a center-point tap, which is the negative leadout of the rectifier output. The two keno-

trons work alternately every half-cycle. Here the rectified-current pulses pass through the load in each half-period and the average rectified current is twice as large as in half-wave rectification. Filters composed of capacitors and chokes are used to smooth the pulsations. A so-called pi-section filter is utilized in the circuit shown. The capacitor C_1 of the filter is connected in parallel to the circuit formed by the choke Dr* and load R. Since the capacitor C_1 possesses a large capacitance, the impedance which it presents to the alternating component of the rectified current is low and the voltage drops of the alternating component across it and, consequently, across the segment of the choke-load circuit parallel to it, are also small. In other words, the capacitance C_1 shunts the load circuit, diverting the alternating component of the rectified current. On the other hand, the choke possesses high impedance and offers little resistance to direct current. For this reason, the alternating component of the voltage applied to the choke-load series circuit is largely lost in the choke and the direct component is applied almost completely to the load. The small load-voltage pulsation is further smoothed out by the capacitor C_2.

The basic parameters of a rectifying circuit are the rectified-current values at a given rectified voltage U. The following method, which is based on the approximate theory of rectifier design, may be used to convert from these circuit parameters to kenotron parameters.

Let us denote the ratio of the rectified voltage to the amplitude of the alternating transformer voltage by η:

$$\eta = \frac{U}{U_{m\text{тр}}}.\text{**} \tag{5-37}$$

*[Др = Dr = Drossel = choke.]
**[$U_{m\text{тр}} = U_{m\text{Tr}} = U_m$ Transformator $= U_m$ Transformer.]

The quantity η is called the rectification efficiency and $\eta = \cos\theta$ when the capacitance of the filter is sufficiently great. In a half-wave circuit, the quantity $U_{m\ Tr}$ is equal to the voltage amplitude in the transformer secondary winding, and in a full-wave circuit to half the voltage amplitude of the whole secondary winding.

The ratio of the plate-current amplitude to the rectified current is a function of the coefficient η:

$$\frac{I_{am}}{\bar{I}} = n = f(\eta). \tag{5-38}$$

The curve of $n = f(\eta)$ is shown in Fig. 5-22. In a full-wave rectifier circuit, the rectified current at one plate is equal to half the total rectified current.

Having determined the value of n, we can find the plate-current amplitude $I_{am} = n\bar{I}$. Obviously, the plate current reaches the amplitude value at the greatest plate voltage. The magnitude of the plate voltage amplitude may be determined as the difference between the transformer voltage amplitude and the filter-capacitor voltage. It is assumed in the calculation that the capacitor's capacitance is so large that the voltage across it remains constant and equal to $\bar{U}$; then

$$U_{am} = U_{mTp} - \bar{U}. \tag{5-39}$$

The tube may be designed on the basis of the values found for I_{am} and U_{am}.

b) Diode detection. Radio-telephone transmission is accomplished by sending into space modulated high-frequency oscillations, i.e., oscillations whose amplitude or frequency varies periodically at a considerably lower (audio) frequency. If the amplitude of the high-frequency oscillations is varied, the modulation is called amplitude modulation (Fig. 5-23a).

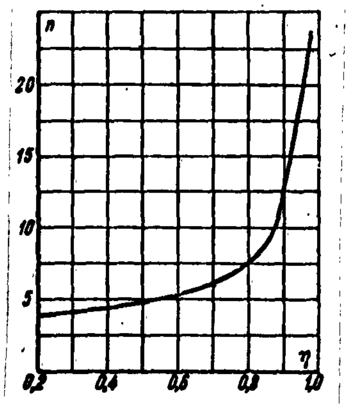

Fig. 5-22. Curve of function $n = f(\eta)$.

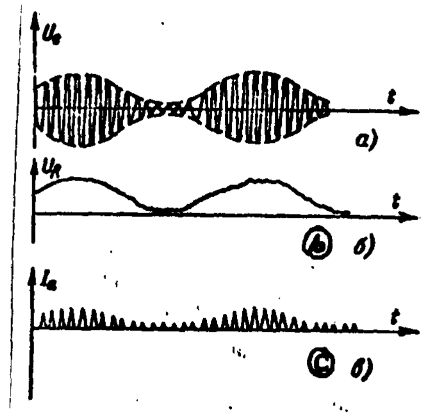

Fig. 5-23. Curves of modulated signal (a), voltage across detector load (b), and diode plate current (c).

In the receiver, it is necessary to separate the low-frequency modulating oscillations which are delivered to the output of the device after amplification. The process of separating the modulating oscillations is called detection.

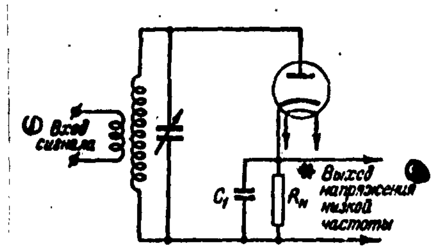

Fig. 5-24. Circuit of diode detector. 1) Signal input; 2) Low-frequency output.

Figure 5-24 shows the simplest diode-detector circuit, which is a half-wave rectifier with an output capacitance. An oscillatory network tuned to the frequency of the signal being received serves as the input of the circuit. A diode and a load shunted by the capacitor C_1 are connected in series with the network. The capacitance C_1 and the load resistance are selected in such a way that the capacitor's discharge time across the load will be much longer than the period of the high-frequency signal oscillations and shorter

*[$R_н = R_n = R_{nagruzka} = R_{load}$.]

than the period of the low-frequency modulating oscillations. Then the capacitor smooths out the high-frequency pulsations of the rectified voltage but cannot smooth out the low frequency tube-current changes. As a result, a voltage drop which varies with a frequency equal to the frequency of the modulating oscillations forms across the load as shown in Fig. 5-23b.

Since the amplitude of the signal voltage is usually small, diodes intended for work in detection circuits work at low plate voltages and, consequently, with small plate currents.

Chapter Six

THE THREE-ELECTRODE TUBE — THE TRIODE

6-1. THE ARRANGEMENT AND PRINCIPLE OF OPERATION OF THE TRIODE

An electron tube in which there is a third electrode — a grid — placed in the space between the plate and the cathode is called a triode. This electrode can be made in the form of a wire mesh or in the form of a helix surrounding the cathode, or it can be made up of a series of rods parallel to the cathode, etc. Figure 6-1 shows the simplest designs for triodes using directly and indirectly heated cathodes; it also shows their schematic symbols.

If the grid has a certain potential with respect to the cathode, it will influence the potential distribution in the space between the anode and the cathode; and, consequently, it will control the number of electrons that leave the space charge. If the grid is negative with respect to the cathode, it will have a retarding action on the electrons leaving the cathode, thus decreasing the plate current. When the grid has a positive potential, it helps the plate, as it were, to attract electrons from the space charge surrounding the cathode. In this case, a part of the electrons fall directly on to the grid, creating a grid current; but the greater part of the electrons reach the plate, and the plate current is increased. Consequently, the grid in a triode is a control electrode, and by changing its potential, the magnitude of the plate current can be varied over wide ranges.

In considering the electric field in the triode, we shall select a plane system analogous to the system that we selected for our consideration of the field in the diode. In the space between the plate and the cathode we shall place a grid in the form of parallel rods (turns) positioned in the plane parallel to the planes of the

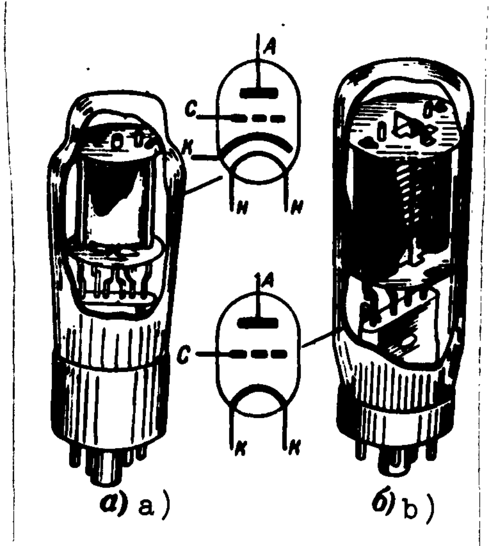

Fig. 6-1. Examples of designs for triodes using an indirectly heated cathode (a) and a directly heated cathode (b).

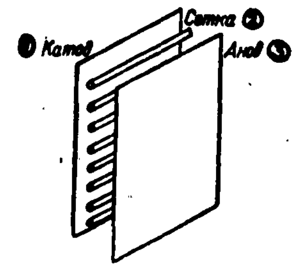

Fig. 6-2. Pictorial representation of a plane triode system. 1) Heater; 2) Grid; 3) Plate.

plate and the cathode, as shown in Fig. 6-2.

The electric field in the interelectrode space of the triode is determined by the common action of the potentials of the plate and the grid. In the plane of the turns of the grid, the potential changes periodically in a direction perpendicular to the turns; the field in the triode is nonhomogeneous.

Figure 6-3a shows a picture of the electric field in a plane triode when the grid has a negative potential and the plate has a positive potential. The electrons from the space charge at the cathode are accelerated by this form of field; and, passing through the interturn space of the grid, fall upon the plate. During this process, almost the whole distance between the turns will permit the passage of electrons, since, except for those sections immediately adjacent to the turns, the region has a positive potential. It is evident that the magnitude of this potential, which is dependent on the potentials on the plate and the grid turns and on the mutual positions of the electrodes, will in turn determine the magnitude of the current taken off

the cathode.

Let us construct the graph of the potential variation in two directions; the first, from the cathode to the plate through the middle of the interturn space; and the second, from the cathode to the plate through the turns of the grid. In constructing the graph, we will, as before, consider the potential of the cathode to be zero. The curves are shown in Fig. 6-3b (the potential minimum at the cathode formed by the space charge is not indicated on the curves).

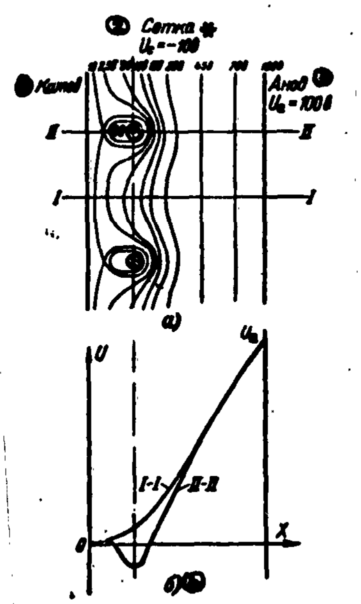

Fig. 6-3. Representation of the electric field in a triode and distribution of potentials when the grid has a negative potential and the space between the turns has a positive potential. 1) Cathode; 2) Grid; 3) Plate.

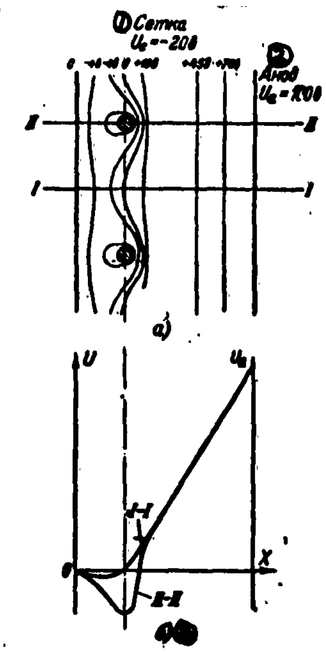

Fig. 6-4. Representation of the electric field in a triode when the space between the turns has a negative potential. 1) Grid; 2) Plate.

The electrons whose motion is directed through the middle of the interturn space, as is evident from the graph, enter a accelerating field and move directly toward the plate. The electrons whose motion is directed toward a turn experience a repulsive force, and part of them return to the cathode. Some of these electrons flow

*[в = v = vol't = volt.]

around the turn, leave the grid, and move through the part of the interturn space which is at a positive potential.

If the potential of the grid is lowered, it is possible to obtain a potential distribution such that the potential in the interturn space will become negative. A representation of such a field is shown in Fig. 6-4. In this case, all the electrons of the space charge that surrounds the cathode are in a retarding field and are unable to move through the grid. The tube is cut off.

In the case where the grid is charged positively, the space between the turns is at a higher potential and a larger current is taken off the cathode. Figure 6-5 shows a representation of the electric field in a triode when the grid is positive.

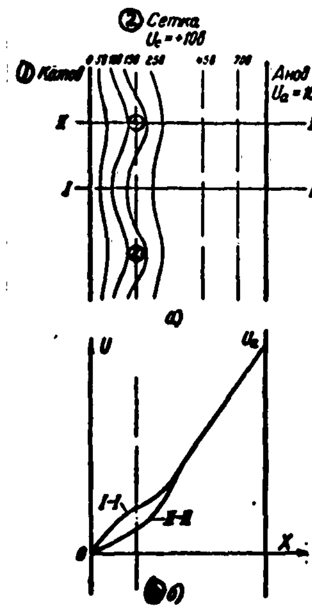

Fig. 6-5. Representation of the electric field in the triode and the potential distribution when the grid has a positive potential.
1) Cathode; 2) Grid; 3) Plate.

In comparing the representations of the electric field for various values of grid potential, it is evident that a variation in the grid potential has a powerful influence on the field in the space

between the cathode and the grid. The electric field in the region between the grid and the plate is changed little by a variation in grid potential.

6-2. THE EQUIVALENT DIODE VOLTAGE OF THE TRIODE

The magnitude of the current taken off the cathode of the triode can be determined if the potential distribution in the plane of the grid is known.

However, since the potential at different points of the grid plane is not the same, such a calculation is very complicated. Therefore, in calculating the currents in the triode, it is expedient to replace this variable grid potential by the averaged potential, constant over the whole surface of the grid, whose action is equivalent to the action of the nonhomogeneous potential of the grid plane. In other words, the common action of the plate and grid of the triode on the cathode may, for purposes of calculation, be replaced by the action of a plate placed at the location of the grid of the triode and having a plate current such that the plate current of the diode thus obtained is equal to the current leaving the cathode in the triode.

This method of calculating the currents in a triode is called the reduction of the triode to the equivalent diode.

The plate potential of the equivalent diode is called the equivalent (resultant) diode voltage. If this voltage is expressed by means of the known magnitudes of the plate and grid potentials, it is possible to calculate the currents of the triode.

In the determination, the electrostatic action of the plate of the equivalent diode on the cathode must be the same as the resultant electrostatic influence of the plate and the grid of the triode on the cathode. The electrostatic influence of one electrode

on another is determined by the charge induced on the latter electrode. The magnitude of this charge is equal to the product of the capacitance between the electrodes and the potential difference between them. Let us denote the capacitance between the plate and the cathode of the triode by $C_{a.k.}$, the capacitance between the grid and the cathode by $C_{c.k.}$, and the capacitance between the plate of the equivalent diode and the cathode by C (Fig. 6-6).

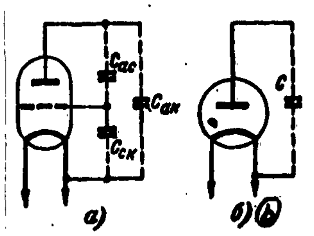

Fig. 6-6. Diagram for the determination of the equivalent diode voltage of the triode. a) Triode capacitances; b) Equivalent diode.

Then the charge induced by the plate of the triode on the cathode is equal to:

$$q_1 = C_{a.k}U_a; \tag{6-1}$$

the charge induced by the grid of the triode on the cathode:

$$q_2 = C_{c.k}U_c; \tag{6-2}$$

the charge induced by the plate of the equivalent diode on the cathode:

$$Q = CU_д,* \tag{6-3}$$

where U_a is the plate potential;

U_c is the grid potential; and

U_d is the equivalent (resultant) diode voltage.

Since the action of the plate of the equivalent diode is equal by definition to the total action of the plate and grid of the triode, we may write:

$$Q = q_1 + q_2. \tag{6-4}$$

*[$U_Д$ = U_d = $U_{deystvuyushchiy}$ = $U_{equivalent}$.]

Substituting the values of the magnitudes of the charges, we obtain:

$$CU_{\text{д}}=C_{\text{с.к}}U_{\text{с}}+C_{\text{а.к}}U_{\text{а}}. \tag{6-5}$$

As an approximation, we may take that $C = C_{c.k} + C_{a.k}$. Then

$$U_{\text{д}}=\frac{C_{\text{с.к}}U_{\text{с}}+C_{\text{а.к}}U_{\text{а}}}{C_{\text{с.к}}+C_{\text{а.к}}}. \tag{6-6}$$

Dividing the numerator and denominator of the righthand side of equation (6-6) by $C_{c.k}$ and writing

$$\frac{C_{\text{а.к}}}{C_{\text{с.к}}}=D,$$

we finally obtain:

$$U_{\text{д}}=\frac{U_{\text{с}}+DU_{\text{а}}}{1+D}. \tag{6-7}$$

Expression (6-7) determines the magnitude of the equivalent diode voltage of the triode. The quantity D is called the penetrance of the triode. Since the penetrance is a ratio between the plate-to-cathode capacitance and the grid-to-cathode capacitance, it is dependent only on the geometry of the tube electrodes. The grid is closer to the cathode than the plate is, and it partially screens the plate from the cathode. Therefore, the plate-to-cathode capacitance is always less than the grid-to-cathode capacitance, and, consequently, the penetrance is less than unity. If $D \ll 1$, it may be neglected in the denominator of Eq. (6-7). A simpler expression is thus obtained for the equivalent diode voltage, and this expression is often used in approximate calculations:

$$U_{\text{д}}=U_{\text{с}}+DU_{\text{а}}. \tag{6-8}$$

The physical significance of the penetrance consists in the fact that it characterizes the degree of penetration of the plate field toward the cathode; and, consequently, it determines the amount of participation of the plate potential in the equivalent diode

voltage of the triode.

By means of the known magnitude of the equivalent diode voltage, it is possible to calculate the currents in the triode. Since the current coming from the cathode of the triode is equal by definition to the plate current of the equivalent diode, the former current may be determined by the three-halves law by substituting the magnitude of the equivalent diode voltage into the formula in place of the magnitude of the plate voltage. Then the magnitude of the current coming from the cathode of the triode may be determined according to the formula

$$I_к = 2{,}33 \cdot 10^{-6} \frac{F_{с.эф}}{r_c^2 \beta^2} U_д^{\frac{3}{2}}. * \tag{6-9}$$

The geometrical coefficient in this formula refers to the dimensions of the plate of the equivalent diode. Therefore, the distance from the plate to the cathode should be taken equal to the grid-to-cathode distance in a triode of plane construction and equal to the radius of the grid in a triode of cylindrical construction. The coefficient β^2 is also a function of the ratio of the grid radius to the cathode radius.

The effective surface for the plate that is positioned in place of the grid is calculated according to the same rules as in the diode.

If the grid has a negative potential, there is no grid current. Then equation (6-9) will be the equation for the plate current of the triode:

$$I_a = 2{,}33 \cdot 10^{-6} \frac{F_{с.эф}}{r_c^2 \beta^2} \left(\frac{U_c + DU_a}{1 + D} \right)^{\frac{3}{2}}. \tag{6-10}$$

When the grid has a negative potential, the cathode current

*[$F_{с.эф.}$ = $F_{s.ef.}$ = $F_{setka.\ effektivnyy}$ = $F_{grid,effective.}$]

I_k is distributed between the grid and the plate of the tube. Denoting the ratio of the currents by

$$\frac{I_a}{I_c} = k,$$

where $\underline{k}$ is the current-distribution factor, and knowing that $I_a + I_c = I_k$, we obtain:

$$I_a = \frac{k}{k+1} I_k. \tag{6-11}$$

Then the equation for the plate current of the triode when the grid has a positive potential takes the following form:

$$I_a = \frac{k}{k+1} \cdot 2{,}33 \cdot 10^{-6} \frac{F_{c.\text{эф}}}{r_c^2 \beta^2} \left(\frac{U_c + DU_a}{1+D} \right)^{\frac{3}{2}}. \tag{6-12}$$

The same corrections that were introduced into the equation for the plate current of the diode are valid for the equation of the plate current in the triode. When the magnitude of the equivalent diode voltage is small, the contact-potential difference between the grid and the cathode must be taken into account.

6-3. THE STATIC CHARACTERISTICS OF THE TRIODE

As was established, the plate current of the triode is a function of two voltages: the plate voltage and the grid voltage. The dependence of the currents in the triode on the voltages in the plate and the grid may be investigated by means of the circuit shown in Fig. 6-7. In this circuit, the power-supply circuits for plate and heater are analogous to those used in investigating the diode. The grid circuit is fed from a power supply E_c by means of a potentiometer and a knife switch, which permits the polarity of the voltage on the grid to be changed. In order to simplify the schematic, the filament circuit has not been shown.

There are two different basic kinds of static triode characteristics: plate characteristics, which represent the dependence of

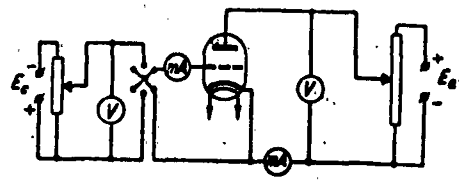

Fig. 6-7. Circuit for investigation of triode.

the plate current of the triode on the plate voltage when the grid voltage is held constant; and plate-grid characteristics which represent the dependence of the plate current on the grid voltage when the plate voltage is held constant. When the characteristics are being taken, a constant filament voltage must be maintained.

In order to plot the plate characteristics, a preset constant grid voltage is established. After this, the plate voltage is gradually increased from zero. If the grid is at a negative potential, there will be plate current only when the equivalent diode voltage becomes positive – i.e., when the plate voltage becomes large enough to overcome the action of the negative grid. Before this happens, while the equivalent diode voltage is still negative, current will not pass through the tube since the tube is cut off. According to the theory, the point at which the plate characteristic begins corresponds to the plate voltage at which the equivalent diode voltage is equal to zero:

$$U_c + DU_{a0} = 0, \tag{6-13}$$

i.e., the magnitude of the plate current corresponding to the beginning of the characteristic is equal to:

$$U_{a0} = -\frac{U_c}{D}. \tag{6-14}$$

A characteristic $I_a = f(U_a)$ for constant grid voltage is shown in Fig. 6-8.

When one plate characteristic has been plotted at some con-

stant value U_c, another voltage may be placed on the grid and a second plate characteristic may be plotted, etc. The group of plate characteristics plotted for various grid voltages is called a family of plate characteristics.

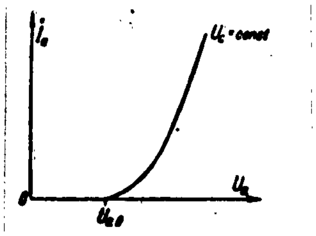

Fig. 6-8. Plate characteristic of a triode.

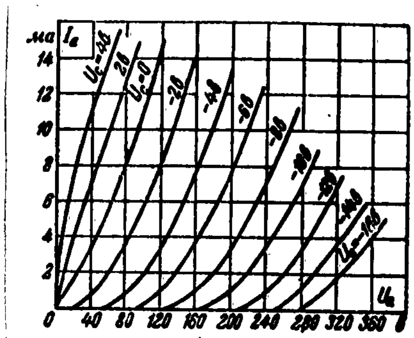

Fig. 6-9. Family of plate characteristics for a type 6N8S triode.

Figure 6-9 shows a family of plate characteristics for a 6N8S [6F8] triode. It is evident from Eq. (6-14) that lowering the grid voltage moves the beginning of the characteristic to the right and raising the grid voltage moves the characteristic to the left. The plate characteristic corresponding to $U_c = 0$ begins at the origin. Characteristics plotted for positive grid voltages also begin at the origin; the greater the magnitude of the grid voltage, the higher the characteristics go.

Plate-grid characteristics are plotted at constant plate voltage.

The point at which the plate-grid characteristic begins corresponds to the grid voltage at which the action of the plate voltage fully compensates for the retarding effect of the negative grid on the electrons. At this point, the plate current is equal to zero; and, consequently, the equivalent diode voltage is also equal to zero. The magnitude of the voltage U_{c0} at which the plate current of the triode is equal to zero is determined from the equation

$$U_{c0}+DU_a=0, \tag{6-15}$$

from which

$$U_{c0}=-DU_a. \tag{6-16}$$

Since the voltage on the plate is always positive, the plate-grid characteristics of the triode always begin at points where the grid voltage is negative.

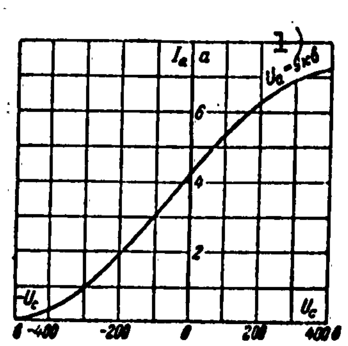

Fig. 6-10. Plate-grid characteristic of the GU-89A triode. 1) kv.

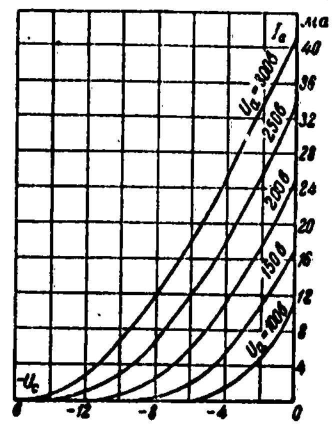

Fig. 6-11. Family of plate-grid characteristics of the 6S5S triode.

The characteristic $I_a = f(U_c)$ is shown in Fig. 6-10. The rising section of the characteristic in the region of negative grid voltages usually approximates a straight line. When one passes into the region of positive grid voltages, the growth of the current is retarded and the characteristic becomes curved. This happens because, in the positive region, a part of the cathode current goes to the grid; as a result of which, the plate current increases more slowly with a growth in the grid voltage.

Anode-plate characteristics may be plotted at various values of plate voltage. The group of characteristics plotted at several values of plate voltage makes up a family of plate-grid triode characteristics. Figure 6-11 shows a family of such characteristics for the 6S5S. It is evident from equation (6-16) that increasing the plate

voltage moves the plate-grid characteristics to the left.

At a given value of plate voltage, the position of the plate-grid characteristic is determined by the penetrance of the tube. The larger the penetrance of the tube, the further to the left the characteristics will be. This is explained by the fact that the larger the penetrance, the less the influence of the plate field on the cathode is weakened; and, consequently, the stronger the retarding effect of the grid must be to cut off the current through the tube. The value of the grid voltage U_{c0} for which the current through the tube is zero is called the grid base or the point of cutoff of the plate voltage. The larger the absolute value of the grid base, the larger is the penetrance of the triode.

6-4. THE STATIC PARAMETERS OF THE TRIODE

The operating conditions of a three-electrode tube are determined by three quantities: the plate voltage, the grid voltage, and the magnitude of the plate current. The connection between changes in two of these quantities when the third is held constant is determined by the static parameters of the triode.

a) The mutual conductance (control-grid – plate transconductance). Since the grid is the control electrode of the triode, it is important to know the extent of its influence on the plate current. When U_a is constant, a definite value of plate current I_a' corresponds to a given value of grid voltage U_c' (Fig. 6-12). Increasing the voltage on the grid to the value U_c'' causes an increase in the plate current to the value I_a''. The ratio of the increment in the plate current $I_a'' - I_a'$ to the increment in the grid voltage $U_c'' - U_c'$ that causes it is called the mutual conductance (control-grid – plate transconductance) of the triode:

$$S=\frac{I_a''-I_a'}{U_c''-U_c'}=\frac{\Delta I_a}{\Delta U_c}\ [\text{ма/в}]. \quad [\text{ma/v}]. \tag{6-17}$$

Expression (6-17) is the definition of the average slope in the interval of grid voltages from U_c^{*} to U_c''. Consequently, the mutual conductance shows how much the plate current of the triode changes when the grid voltage is changed by 1 volt; in other words, how steeply the plate-grid characteristic rises.

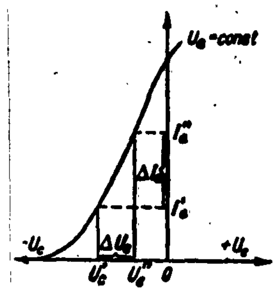

Fig. 6-12. Derivation of triode transconductance.

The changes in the grid voltage may be quite small in practice. Therefore, it is essential to determine the mutual conductance for sufficiently small changes in the grid voltage – i.e., as $\Delta U_c \to 0$.

Then

$$S=\lim_{\Delta U_c \to 0}\frac{\Delta I_a}{\Delta U_c}=\frac{dI_a}{dU_c} \tag{6-18}$$

when U_a = const – i.e., the mutual conductance is the derivative of the plate current with respect to the grid voltage. Substituting the value of I_a obtained from (6-10) and differentiating, we obtain:

$$S=\frac{3}{2}\cdot 2{,}33\cdot 10^{-6}\,\frac{F_{c.эф}}{r_c^{2}\beta^{2}}(U_c+DU_a)^{\frac{1}{2}} \tag{6-19}$$

or

$$S=3{,}5\cdot 10^{-6}\,\frac{F_{c.эф}}{r_c^{2}\beta^{2}}\,U_a^{\frac{1}{2}}. \tag{6-19a}$$

Consequently, the mutual conductance is porportional to the square root of the magnitude of the equivalent diode voltage. Since

Eq. (6-10) is valid only when the grid voltage is negative, expression (6-19a) for the mutual conductance is also only valid in the region $U_c < 0$. When the grid is positive, the mutual conductance decreases owing to the appearance of a grid current. In order to determine the mutual conductance when the grid is positive, the multiplier $k/(k + 1)$ must be introduced into Eq. (6-19a), where k is the current distribution factor. Then the expression for the mutual conductance takes the following form:

$$S = \frac{k}{k+1} 3{,}5 \cdot 10^{-6} \frac{F_{c.э\phi}}{r_c^2 \beta^2} U_a^{\frac{1}{2}}. \qquad (6\text{-}20)$$

An analysis of Eqs. (6-19a) and (6-20) shows that, at given plate and grid voltages, the mutual conductance becomes larger if the effective surface of the grid is made larger or the grid-cathode spacing is made smaller. The effective surface of the grid, and, consequently, the mutual conductance, can be made larger by increasing the length of the grid; but this increases the length of the whole tube construction. Increasing the length of tube elements, including the cathode, leads to an undesirable increase in the amount of power required for the filament; furthermore, the mechanical stability required for the tube construction limits the amount of lengthening that can be done. A better method of increasing the mutual conductance of the triode is to decrease the grid-cathode spacing, since the mutual conductance is inversely proportional to the square of this spacing. In modern directly heated triodes, the minimum value of this spacing amounts to 100 to 120 microns; and in indirectly heated tubes, to 60 to 70 microns. In certain special designs of indirectly heated triodes, this spacing only amounts to 15 to 20 microns.

If the grid-cathode spacing is increased and the other dimensions of the grid are held constant, the mutual conductance will rise quickly at first, but then it will start to decrease. The explanation

for this is that, when the spacing between the grid and the cathode is very small, the influence of the grid does not extend to the whole cathode. Individual sections of the cathode are under the direct influence of the field of the plate (Fig. 6-13). The value of the current from these sections depends little on the grid potential. This phenomenon is called the "island" effect. In order to avoid the "island" effect when the grid-cathode spacing is decreased, the spacing between the windings of the grid must be decreased; this latter spacing is called the grid pitch. The grid pitch should be made less than twice the grid-cathode spacing.

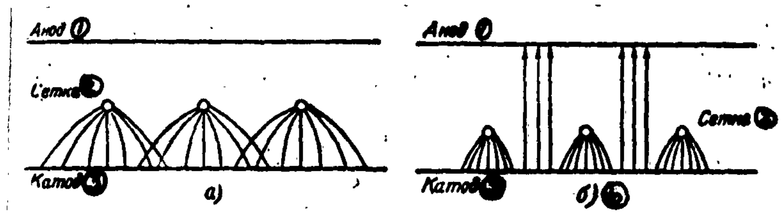

Fig. 6-13. "Island" effect in the triode. a) The whole cathode is under the influence of the grid; b) part of the cathode (islands) are not controlled by the grid. 1) Plate; 2) Grid; 3) Cathode.

b) The static amplification factor of the triode. As has already been pointed out, the plate current of a triode can change when either the plate voltage or the grid voltage is changed. A change in the grid voltage has a greater influence on the value of the plate current than the same change in the plate voltage, since the grid is nearer to the cathode and screens the plate from the cathode. Consequently, the same change in the plate current may be obtained by different variations in the plate and grid voltages. If the voltage on the grid is increased by an amount ΔU_c, the plate voltage must be decreased in order to keep the same plate current. It is evident that the magnitude of the change in the plate voltage ΔU_a must be greater than ΔU_c. The absolute value of the ratio of the increment in plate

voltage to the increment in grid voltage when the plate current is held constant is called the amplification factor of the triode:

$$\mu=\left|\frac{\Delta U_a}{\Delta U_c}\right| \text{where } I_a=\text{const.} \tag{6-21}$$

The amplification factor shows by what factor the influence of the grid voltage on the plate current is greater than that of the plate voltage.

The amplification factor is measured when the plate current is constant; and, consequently, the changes in the plate and grid voltages have different signs. A decrease in the plate voltage corresponds to an increase in the grid voltage, and vice versa. In virtue of this, the ratio of the increments in the plate and grid voltages at constant plate current is always a negative quantity. Therefore, the amplification factor is usually defined as the absolute value of this ratio.

The constancy of the plate current in the measurement of the amplification factor corresponds to the invariance of the equivalent diode voltage when the plate and grid voltages are being simultaneously changed.

Let us take two static plate-grid characteristics plotted at plate voltages of U_a' and $U_a' + \Delta U_a$ (Fig. 6-14). At a constant plate current I_a', a grid voltage U_c' will correspond to a plate voltage U_a'; and a grid voltage $U_c' + \Delta U_c$ will correspond to a plate voltage $U_a' + \Delta U_a$. Since the plate current is constant, the equivalent diode voltages for both combinations of plate and grid voltages are equal to each other:

$$U_c'+DU_a'=(U_c'+\Delta U_c)+D(U_a'+\Delta U_a). \tag{6-22}$$

Opening the parentheses and grouping the terms with D on the right-hand side of the equation, we obtain:

$$-\Delta U_c=D\Delta U_a. \tag{6-23}$$

from which

$$\left|\frac{\Delta U_a}{\Delta U_c}\right| = \frac{1}{D}. \tag{6-24}$$

The lefthand side of Eq. (6-24) is the amplification factor. Consequently,

$$\mu = \frac{1}{D}. \tag{6-25}$$

The amplification factor is a quantity that is inversely proportional to the penetrance, and, consequently, is dependent only on the geometry of the electrodes of the tube. Since

$$D = \frac{C_{a.k}}{C_{c.k}},$$

we have

$$\mu = \frac{C_{c.k}}{C_{a.k}} \tag{6-26}$$

and the problem of determining the dependence of the amplification factor on the geometry of the electrodes is reduced to the calculation of the interelectrode capacitances of the triode. However, it is difficult to calculate these capacitances accurately even in the simplest tube designs. It is evident that the amplification factor will be greater, the smaller the grid pitch and the larger the diameter of the wires of the windings; since, when these two things are done, the grid-cathode capacitance is increased, and the shielding of the plate from the cathode is increased, which decreases the plate-cathode capacitance. Moving the plate further away from the grid also increases the amplification factor because the plate-cathode capacitance is lowered.

c) The dynamic plate resistance (internal resistance) of the triode. The connection between the changes in the plate voltage and the plate current when the grid voltage is held constant is determined by a third parameter of the triode — the dynamic plate resistance (internal resistance). Let us consider the static plate char-

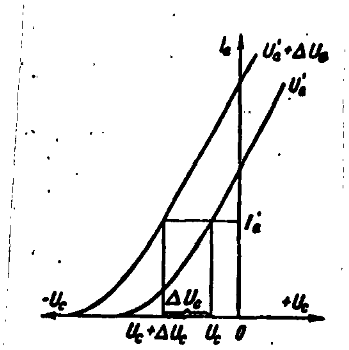

Fig. 6-14. Computing the amplification factor of the triode.

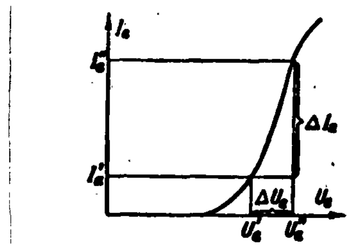

Fig. 6-15. Computing the internal resistance of the triode.

acteristic of a triode (Fig. 6-15). A plate voltage U_a' corresponds to a plate current value I_a'. In order to change the plate current to a value I_a'', the plate voltage must be made equal to U_a''. The ratio of the changes in the plate voltage to the change in the plate current when the grid voltage is held constant is called the dynamic plate resistance of the triode:

$$R_i = \frac{U_a'' - U_a'}{I_a'' - I_a'} = \frac{\Delta U_a}{\Delta I_a}. \tag{6-27}$$

When, as is usual, the plate current is measured in milliamperes and the plate voltage in volts, the unit of dynamic plate resistance is the kilohm.

Thus, the dynamic plate resistance determines by how many volts the magnitude of the plate voltage must be changed to produce a change in the plate current of 1 ma, when the grid voltage is held constant.

Using the circuit for the investigation of the triode, all three triode parameters may be measured by the following method. Nominal voltage values are established on the tube electrodes, and the value of the plate current is measured at these voltages. Then the voltage on the grid is changed while the plate voltage is held constant, and the value of the plate current is measured. For the

third reading, the grid voltage is held at a constant value equal to its value at the second reading, and the plate voltage is varied until the magnitude of the plate current becomes equal to its value at the nominal plate and grid voltages – i.e., to the value of the plate current at the first reading. The results of the measurements may be collected into a table:

№ отсчета[1]	U_a	U_c	I_a
1	U'_a	U'_c	I'_a
2	U'_a	U''_c	I''_a
3	U''_a	U''_c	I'_a

1) No. of reading.

By the results of the first and second readings, which correspond to the change in the plate current when the grid voltage is changed and the plate voltage is held constant, the mutual conductance of the triode may be determined:

$$S = \frac{I''_a - I'_a}{U''_c - U'_c}. \qquad (6\text{-}28)$$

The dynamic plate resistance may be determined from the second and third readings, since these readings correspond to a change in the plate voltage when the grid voltage is held constant:

$$R_i = \frac{U''_a - U'_a}{I'_a - I''_a}. \qquad (6\text{-}29)$$

The first and third readings make it possible to determine the amplification factor of the triode – i.e., the ratio of the magnitudes of simultaneous changes in the plate and grid voltages when the plate current is held at a constant value equal to the nominal value:

$$\mu = \left|\frac{U''_a - U'_a}{U'_c - U''_c}\right|. \qquad (6\text{-}30)$$

Since $\mu = 1/D$,

$$D = \frac{U'_c - U''_c}{U''_a - U'_a}. \qquad (6\text{-}31)$$

A connection can be established between the basic parameters from the expressions that have been obtained for them. Let us multiply the values for the mutual conductance, the dynamic plate resistance, and the penetrance obtained by the three-readings method:

$$DSR_i = \frac{U_c' - U_c''}{U_a'' - U_a'} \frac{I_a'' - I_a'}{U_c'' - U_c'} \frac{U_a'' - U_a'}{I_a' - I_a''}. \qquad (6\text{-}32)$$

By cancellation, we obtain the equation that connects the three basic parameters of the triode:

$$DSR_i = 1. \qquad (6\text{-}33)$$

Equation (6-33) is called the internal equation of the triode. Substituting the quantity $1/\mu$ into the equation in place of the penetrance, we obtain another, more common form of this equation

$$SR_i = \mu. \qquad (6\text{-}34)$$

The internal equation of the triode is valid for any operating conditions of the triode.

The triode parameters can also be determined graphically from a family of static characteristics. For this, two characteristics are selected that have been plotted at voltages close to the nominal voltage. In determining the parameters from the plate static characteristics (Fig. 6-16), a characteristic triangle ABC is constructed so that points A and B correspond to the nominal value of the plate current I_a'. The legs of this triangle are the increments in the plate current ΔI_a and the increment in the plate voltage ΔU_a. The increment in the grid voltage is determined as the difference of the values of the grid voltage known for both characteristics:

$$\Delta U_c = U_c'' - U_c'.$$

Knowing the values of all increments, the parameters of the triode may be computed by formulas (6-17), (6-21), and (6-27).

Analogously, a characteristic triangle may be constructed on

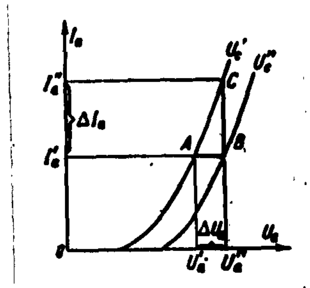

Fig. 6-16. Determination of triode parameters from plate characteristics.

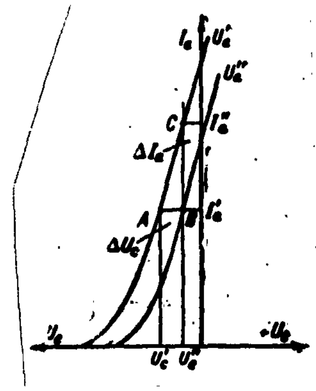

Fig. 6-17. Determination of triode parameters from plate-grid characteristics.

two plate-grid characteristics of the triode (Fig. 6-17). In this case, the legs of the characteristic triangle will be ΔU_c and ΔI_a. The magnitude of the increment in the plate voltage is determined as the difference in the voltages known for each characteristic. The accuracy of the parameter determination by the indicated methods is dependent on the magnitude of the interval of the voltages between readings. It is possible to use the method that is used for the diode in determining the value of the mutual conductance at a given point of the characteristic. The mutual conductance will be the tangent of the angle of inclination of the tangent to the given point of the plate-grid characteristic, computed according to the scale of the characteristic. Analogously, the dynamic plate resistance will be the cotangent of the angle of inclination of the tangent to the plate characteristic.

The methods cited for determining the parameters of tubes may be applied for any triode operating conditions that will permit the selection of operating conditions according to the existing static characteristics.

6-5. THE DISTRIBUTION OF CURRENTS IN THE TRIODE

Many types of triodes operate with positive potentials on the grid. When this is done, the current taken from the cathode is distributed between the plate and the grid of the tube.

The calculation of the current distribution in the triode is based on the theorem of the similarity of electric fields. The essence of this theorem is that, if in any system of charged conductors the potentials of all conductors are changed by the same factor, the form of the electric field does not change; the field formed at the changed potential values will be "similar" to the field at the former conductor potentials.

When applied to the triode, this theorem means that if the potentials of the grid and the plate are simultaneously changed by the same factor, the paths of the electrons do not change, since the form of the lines of force of the electric field remains the same as at the former potentials. It follows from this that the ratio of the magnitude of the plate current to the grid current is not dependent on the absolute values of the plate and grid voltages, but on their ratio. Consequently, the current distribution factor $k = I_a/I_c$ is a function of the ratio U_a/U_c:

$$\frac{I_a}{I_c} = f\left(\frac{U_a}{U_c}\right). \qquad (6\text{-}35)$$

The nature of the distribution of currents in the triode changes according to the magnitudes of the voltages on the plate and grid. If $U_a > U_c$, any electron that goes beyond the grid into the grid-plate space will enter an accelerating field and will reach the plate. In this case, the grid current will be due only to the electrons that fall directly on the turns of the grid. These electrons are intercepted, as it were, by the grid on their way to the plate; and, therefore, such current-distribution conditions are called

direct-interception conditions. Figure 6-18 shows a representation of this type of current distribution.

If the plate voltage becomes less than the voltage on the grid, the electrons that land on the grid include not only directly intercepted electrons, but also those returned from the space charge retarded in the grid plate space. When this happens, the nature of the current distribution changes.

The mechanism of the distribution of the currents between the plate and grid under electron-return conditions is determined by the retarding field in the grid-plate space. In order fully to decelerate the electrons that pass through the grid toward the plate, the magnitude of the potential difference that sets up the retarding field must be equal to or greater than the potential difference that the electrons have passed through in their accelerated motion toward the grid. Let us suppose that the velocity of the electrons that pass through the grid is determined by the grid voltage U_c. Since the plate voltage is always greater than zero, the potential difference between the grid and the plate is always less than the grid voltage. Consequently, if the motion of an electron is normal to the surface of the plate, it will not lose all of its velocity on its way from the grid to the plate and will land on the plate. But if the v ocity of the electron is directed at an angle to the surface of the plate, the velocity component in the direction of the intensity of the retarding field can be small; and, at some distance from the grid, the electron will lose all of its velocity component in the direction of the plate; and, under the influence of the electric field, will return to the grid, as shown in Fig. 6-19.

When $U_a < U_c$, the potential of the points of the interturn space is less than the potential of the grid turns. As a result, an

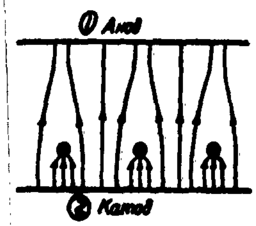

Fig. 6-18. Current distribution under direct interception conditions (the arrows show the paths of the electrons). 1) Plate; 2) Cathode.

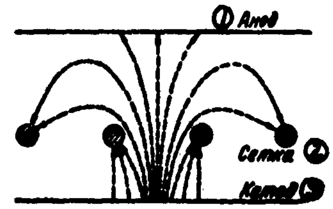

Fig. 6-19. Path of the electrons in a triode under electron-return conditions. 1) Plate; 2) Grid; 3) Cathode.

electric-field component appears that is directed from the turns to the middle of the interturn space. Consequently, the interturn space acts like a diverging lens on the stream of electrons, and the electrons pass through the grid velocities directed at an angle to the surface of the grid. There will be no curvature of path only for the electrons that pass through the grid at the exact center between the turns. The nearer the path of an electron is to a turn, the more its path will be curved, and the less will be its velocity component in the direction of the plate.

When the current density is large enough, the retarded electrons form a space charge of great density in the grid-plate space; the presence of this charge leads to the appearance of a potential minimum in this space. Part of the electrons from the space charge that accumulates in the region of potential minimum are taken by the plate, and the rest return to the grid. Such a retarded space charge is called a fictitious or virtual cathode.

When the plate potential is increased, the number of electrons that return to the grid quickly decreases, since, when the plate potential is increased, the intensity of the retarding electric field is decreased, and, therefore, the dispersion of the electron paths in the interturn space is also decreased.

In practice, electron return ceases to occur not when the plate and grid voltages are equal, but when the plate voltage is still less than the grid voltage. Direct interception conditions have already begun to hold when $U_a > 0.8\ U_c$. The explanation for this is that, when the plate voltage is only a little less than the grid voltage, the electrons do not fully lose their velocity on their way from the grid to the plate.

An investigation of the dependence of the current-distribution factor on the ratio of the plate voltage to the grid voltage under direct interception conditions produces the following expression:

$$\frac{I_a}{I_c} = C_1\sqrt{\frac{U_a}{U_c}}. \tag{6-36}$$

where C_1 is a quantity dependent on the geometry of the tube electrodes.

The greater the grid pitch and the smaller the thickness of the wires of the turns, the larger this quantity will be. Furthermore, the coefficient C_1 is dependent on the ratio of the plate-cathode spacing to the grid-cathode spacing. For a plane electrode system

$$C_1 = \frac{t}{d}\left(\frac{r_c}{r_a}\right)^{\frac{2}{3}}, \tag{6-37}$$

where $\underline{t}$ is the grid pitch;

$\underline{d}$ is the diameter of the wires of the grid turns;

r_a and r_c are the plate-cathode and grid-cathode spacings respectively.

The ratio $d/t = \alpha$ is the fill factor of the grid.

When the electrode system is cylindrical, the factor C_1 has the following value:

$$C_1 = \frac{t}{d}\left(\frac{r_c}{r_a}\right)^{\frac{1}{3}}. \tag{6-38}$$

In formula (6-38), the quantities r_a and r_c denote the radii of the plate and the grid reppectively.

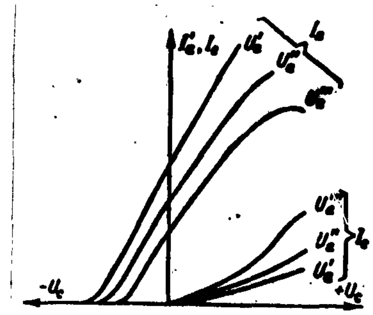

Fig. 6-20. Characteristics of the plate and grid currents in the region of positive grid voltages.

For approximate calculation of the current distribution under electron-return conditions, one may use the expression

$$\frac{I_a}{I_c} = 1.4C_1\left(\frac{U_a}{U_c}\right)^2. \tag{6-39}$$

Thus, in passing into the region of positive grid-voltage values, part of the cathode current goes to the grid, as a result of which the plate current grows more slowly. Figure 6-20 shows typical plate-current and grid-current characteristics plotted at various plate voltages. In the region of positive grid voltages, the growth in the plate current, which determines the mutual conductance, is lowered more on the characteristics which correspond to lower plate voltages, since the part of the current that goes to the grid is greater. Correspondingly, the grid-current characteristics go higher when the plate voltages are lower.

6-6. THE INTERELECTRODE CAPACITANCES OF THE TRIODE

When considering the use of the triode in various circuits, it is essential to know not only the parameters that have been discussed, but also the interelectrode capacitances, especially in high-

frequency circuits — i.e., radio-frequency circuits. Since the capacitive susceptance is proportional to the frequency of the applied voltage, capacitive currents that violate the operating conditions of the tube may arise when the tube is operated at high frequencies.

The interelectrode capacitances of the triode are dependent not only on the geometry of the electrodes themselves, but on the capacitances between the leads and the supports of the electrodes. In conventional triode designs, the capacitances between the leads and the supports make up 50 per cent of the interelectrode capacitances.

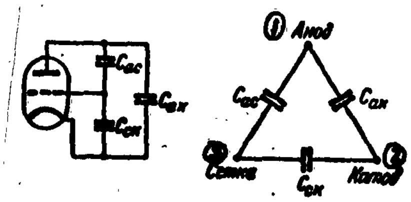

Fig. 6-21. Capacity triangle of the triode. 1) Plate; 2) Cathode; 3) Grid.

The interconnection of the interelectrode capacitances may be represented schematically by means of a capacity triangle (Fig. 6-21), the vertices of which are the tube electrodes; the cathode, the grid, and the plate. It is impossible to make a direct measurement of each of the interelectrode capacitances separately, since when any two electrodes are connected to a measuring bridge to measure the capacitances between them, the measurement will be distorted by the fact that this capacitance is shunted since it has two other series-connected capacitances in parallel with it. Thus, for example, if the plate and grid are connected to a measuring bridge in order to measure the capacitance $C_{a.c}$, the two series-connected capacitances $C_{a.k}$ and $C_{c.k}$ will be connected in parallel with the sought capacitance.

In order to determine the interelectrode capacitances, a

method is used that consists in the consecutive measurement of the capacitance of each electrode with respect to the two other capacitances interconnected. When this is done, the sum of two parallel-connected capacitances is measured in each case. When the plate and grid are connected, it is possible to measure the capacitance of the cathode with respect to the plate and the grid:

$$C_{k}=C_{c.k}+C_{a.k}. \qquad (6-40)$$

Analogously, if the grid and cathode are connected, we obtain:

$$C_{a}=C_{a.k}+C_{a.c} \qquad (6-41)$$

and, if the plate and cathode are connected,

$$C_{c}=C_{c.k}+C_{a.c}. \qquad (6-42)$$

These expressions make up a system of three equations in three unknowns, the sought capacitances: $C_{a.k}$, $C_{a.c}$, $C_{c.k}$. Solving these equations, we obtain for the interelectrode capacitances:

$$C_{a.k}=\frac{C_a+C_k-C_c}{2}; \qquad (6-43)$$

$$C_{a.c}=\frac{C_a+C_c-C_k}{2}; \qquad (6-44)$$

$$C_{c.k}=\frac{C_c+C_k-C_a}{2}. \qquad (6-45)$$

The grid-cathode capacitance $C_{c.k}$ is called the input capacitance of the triode; the plate-grid capacitance $C_{a.k}$, the output capacitance; and the plate-grid capacitance $C_{a.c}$, the transfer capacitance.

6-7. THE DEPENDENCE OF THE PARAMETERS OF THE TRIODE ON THE GEOMETRY OF THE ELECTRODES

The determination of the dependence of the parameters of the triode on the geometry of the electrodes is connected with the determination of the electric fields in the tube. A more or less accurate theoretical formulation of this dependence can be worked out for the

simplest electrode forms, the plane and cylindrical forms. However, the forms of the electrodes of actual tubes are more complicated due to design and technological considerations.

The calculation of the parameters of actual tubes can only be approximate, and the final dimensions of the electrodes are determined experimentally on experimental models.

a) The mutual conductance. Formulas (6-19a) and (6-20) establish a connection between the mutual conductance of the triodes and the dimensions of the electrodes for ideal plane and cylindrical tube designs. These formulas are valid on condition that the whole surface of the cathode is homogeneously covered by the grid — i.e., if the distance to the grid from any point on the surface of the cathode is the same. These formulas also do not take into account the influence of the transverses of grids, and this influence can be quite important when the spacing between the transverses and the cathode is small. In complicated triode systems, the distribution of the space charge strongly influences the value of the mutual conductance; but it is not always possible to calculate this distribution.

When the tube has a cylindrical indirectly-heated cathode and an oval grid (Fig. 6-22), the mutual conductance should be calculated according to the usual formula for a cylindrical system after introducing a coverage factor γ for the cathode into it:

$$S = 3{,}5 \cdot 10^{-6} \gamma \frac{F_{c.эф}}{r_c^2 \beta^2} U_a^{\frac{1}{2}}. \qquad (6\text{-}46)$$

In this formula, the quantity r_c is taken equal to half of the grid span ($r_c = a/2$), and the factor γ varies from 0.4 to 0.8 in different tube designs.

In order to make better use of the cathode surface, modern receiving tubes (including low-power amplifiers) often employ plane

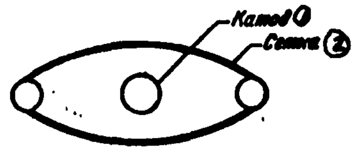

Fig. 6-22. Cylindrical-cathode – oval-grid system. 1) Cathode; 2) Grid.

cathodes with plane grids and oval cathodes with grids whose surfaces are ovals spaced evenly from the cathode surface (Fig. 6-23a and b). With the same aim in mind, use is often made of crimped grids, which give a cathode coverage factor of close to unity (Fig. 6-23c).

When calculating its mutual conductance, the system shown in Fig. 6-23b should be considered as plane.

Since the mutual conductance is dependent on the effective surface of the grid, and on the factor β^2, any necessary increase in the mutual conductance should be accomplished, assuming minimum permissible grid-cathode spacing, by increasing the length of the cathode or decreasing the factor β^2 by increasing the radius of the cathode. In both cases, the surface of the cathode is increased; and, consequently, the power required for heating is also increased. Therefore, in designing receiving tubes, the factor that determines the cathode surface is often not the value of the cathode current, but the value of mutual conductance required.

b) The amplification factor of the triode. Analysis of the electrostatic field in the triode shows that the value of the amplification factor is a function of the interelectrode spacings and the fill factor, which is defined as the ratio of the part of the grid surface that is occupied by the turns to the whole surface.

The formulas that exist for calculating the amplification

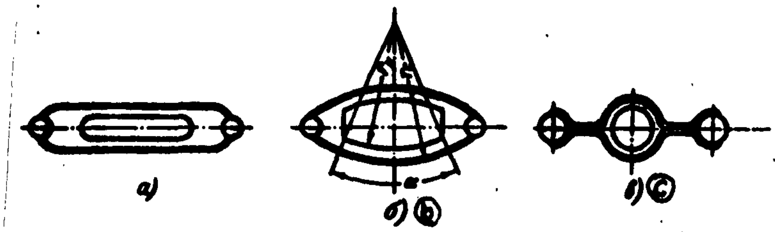

Fig. 6-23. System with improved use of cathode surface. a) Plane system; b) Oval system; c) Crimped-grid system.

factor give sufficiently accurate results only for ideal plane and cylindrical electrode systems. The application of these formulas to actual designs produces errors that are eliminated after the manufacture of experimental tube models by a conversion of the amplification factor actually obtained to the required value.

In calculating the amplification factor of triodes of cylindrical construction, one may employ the formula

$$\mu_{ц} = L_c \frac{\lg \frac{r_a}{r_c}}{f(\alpha)} \cdot * \qquad (6\text{-}47)$$

In this formula, L_c is the so-called active wire length – i.e., the length of turns of wire per 1 cm of grid length. The function $f(\alpha)$ of the fill factor is shown on the graph in Fig. 6-24.

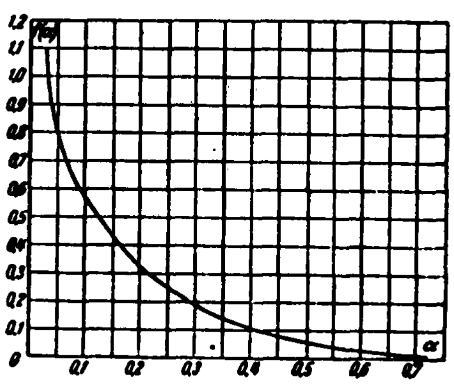

Fig. 6-24. Graph of the function $f(\alpha)$.

The analogous formula for a plane system of triode electrodes has the form:

*[lg = log. $U_{Ц} = U_{Ts} = U_{tsilindricheskiy} = U_{cylindrical}$.]

$$\mu_п = \frac{L_c 2\pi (r_a - r_c)}{f(\alpha)} .* \qquad (6\text{-}48)$$

The formulas for computing the active wire length of grids and the fill factor for various designs are collected in Table 6-1.

TABLE 6-1

(1) Конструкция сетки	(2) Цилиндрическая система		(3) Плоская система	
	α	L_c	α	L_c
(4) Спиральная сетка	$\frac{d_c}{t}$	$\frac{2\pi r_c}{t}$	$\frac{d_c}{t}$	$\frac{1}{t}$
(5) Спиральная сетка с траверсами	$\frac{d_c}{t} + N\frac{d_c}{2\pi r_c} \times \left(1 + \frac{d_т}{t}\right)$	$\frac{2\pi r_c}{t} + N \times \left(1 - \frac{d_т}{t}\right)$	$\frac{d_c}{t}$	$\frac{1}{t}$
(6) Плетеная сетка	$\frac{d_c}{a} \times \left(2 - \frac{d_c}{a}\right)$	$\frac{4\pi r_к}{a} \times \left(1 - \frac{d_c}{2a}\right)$	$\frac{d_c}{a} \times \left(2 - \frac{d_c}{a}\right)$	$\frac{1}{2} \times \left(2 - \frac{d_c}{a}\right)$

1) Grid design; 2) Cylindrical system; 3) Plane system; 4) Helical grid; 5) Helical grid with traverses; 6) Mesh grid.

The symbols employed in the Table for the grid dimensions are explained in Fig. 6-25; N is the number of traverses.

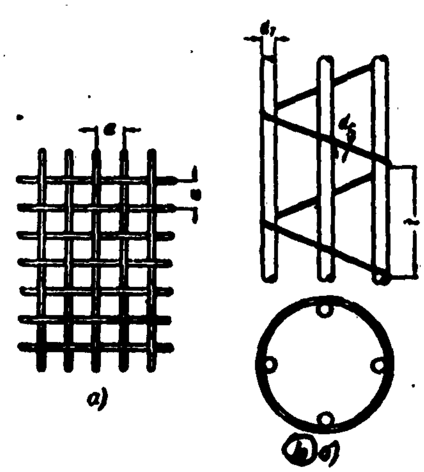

Fig. 6-25. Symbols for grid dimensions. a) Mesh grid; b) Helical grid on traverses.

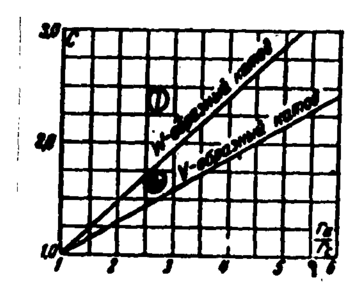

Fig. 6-26. Graph of $C = f(r_a/r_c)$. 1) W-shaped filament; 2) V-shaped filament.

*[$\mu_п = \mu_p = \mu_{ploskiy} = \mu_{plane}$.]

For a plane triode system with a directly heated loop-type filament, good results are given by the proposed formula of Kuzunoz

$$\mu = \frac{C L_c \lg \frac{r_a}{r_c}}{\varphi(\alpha)}, \qquad (6\text{-}49)$$

where C is dependent on the ratio r_a/r_c.

A graph of $C = f(r_a/r_c)$ is shown in Fig. 6-26 for filaments in the form of one or two loops; and Fig. 6-27 shows a graph of the function $\varphi(\alpha)$.

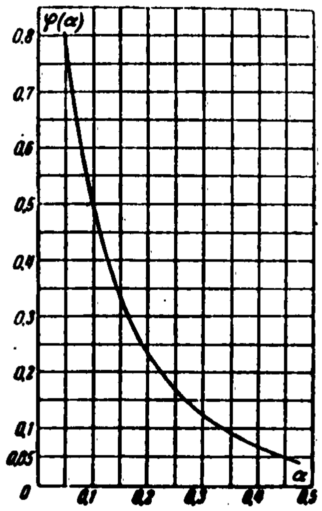

Fig. 6-27. Graph of the function $\varphi(\alpha)$.

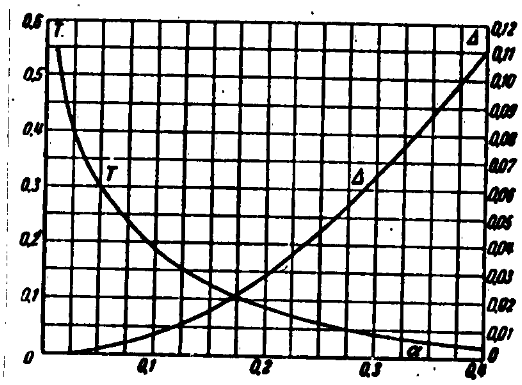

Fig. 6-28. Graph of the functions T and Δ.

Wide use is made of Ollendorf's formula:

$$\mu = \frac{b_2}{tT} - \frac{\Delta}{T}. \qquad (6\text{-}50)$$

where b_2 is the smallest grid-plate spacing;
T and Δ are functions of the fill factor α.

This formula is in good accord with experimental results when the fill factor $\alpha < 0.4$, which covers the majority of triode designs. Figure 6-28 depicts graphs of the functions T and Δ.

Chapter Seven

TRIODE APPLICATIONS

7-1. DYNAMIC OPERATION OF A TRIODE.

In the previous chapter, we considered the operation of a triode under static working conditions, under which the voltages on the plate and grid, determining the operation of the tube, could be measured independently. In this case, triode operation was considered without reference to the external circuit.

In practice, in any circuit in which a triode is used, it is necessary to connect into its plate circuit (sometimes into the cathode circuit) a resistance — the load. Depending upon the circuit, this resistance may be purely resistive, be active, or composite. The load is the first element to receive electrical fluctuations occurring in the tube as a result of the action of the varying voltages applied to the tube grid.

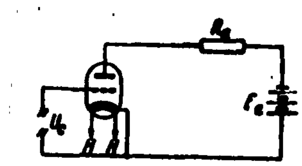

Fig. 7-1. Diagram of triode dynamic operating regime.

The simplest case is that in which a triode operates in a circuit with a resistive load in the plate circuit (Fig. 7-1).

For any value of voltage U_c greater than the cutoff voltage, a current I_a will flow in the plate circuit. This current creates a voltage drop across the load of

$$U_R = I_a R_a. \tag{7-1}$$

According to Kirchhoff's law, with a closed circuit, it is possible to write an expression for the plate circuit

$$U_a + I_a R_a = E_a, \qquad (7-2)$$

i.e., the sum of the voltage drops across the tube and the load will equal the plate supply voltage E_a. From this it follows that

$$U_a = E_a - I_a R_a. \qquad (7-3)$$

If the voltage on the triode grid changes, the change in plate current thus caused will in turn lead to a change in the voltage drop across the load. Since the plate supply voltage E_a remains constant, then in accordance with Eq. (7-3), the plate voltage U_a should change. Consequently, when a triode is operating with a load connected into the plate circuit, a change in grid voltage will cause a change in plate voltage. This operating regime of a triode is called the dynamic (or working) regime.

It should be noted that a change in grid voltage corresponds to a change in plate voltage in the opposite direction. For example, where the grid voltage rises, the plate current will increase and, consequently, the voltage drop across the load will increase. With the increase in voltage drop across the load, the plate voltage drops, as is clear from Eq. (7-3). Conversely, a drop in grid voltage leads to a rise in plate voltage.

Clearly, under dynamic operating conditions, a change in grid voltage leads to a smaller change in plate current than occurs under static conditions, since it is accompanied by a change in plate voltage that opposes the change in plate current.

7-2. DYNAMIC CHARACTERISTICS.

The connection between plate current and plate voltage under

dynamic conditions is defined by an expression derived from Eq. (7-3):

$$I_a = \frac{E_a}{R_a} - \frac{1}{R_a} U_a. \qquad (7\text{-}4)$$

It is clear from Eq. (7-4) that the plate current of a triode, under dynamic conditions, is proportional to the plate current itself. This relationship may be represented graphically by a straight line. Such a straight line is called a dynamic plate characteristic curve (or load line).

In order to construct a dynamic plate curve, it is necessary to determine the position of two points on this line. Setting I_a equal to zero in Eq. (7-4), we obtain

$$\frac{E_a}{R_a} - \frac{1}{R_a} U_{a1} = 0, \qquad (7\text{-}5)$$

whence

$$U_{a1} = E_a. \qquad (7\text{-}5a)$$

This point on the curve corresponds to a cut-off tube. Since in this case the voltage drop across the load is zero, the entire plate-supply voltage appears across the tube.

A second point on the curve may be obtained by setting the plate voltage U_a equal to zero in Eq. (7-4):

$$I_{a1} = \frac{E_a}{R_a}. \qquad (7\text{-}6)$$

Plotting $U_a = E_a$ along the voltage axis, and $I_a = E_a/R_a$ along the current axis, and joining these points by a straight line, we obtain a dynamic plate curve that is called the load line (Fig. 7-2).

The slope of the curve is determined by the value of the plate load. From the curve plotted in Fig. 7-2, it is clear that

$$\operatorname{tg} \alpha = \frac{\frac{E_a}{R_a}}{E_a} = \frac{1}{R_a}. \qquad (7\text{-}7)$$

By changing the load resistance, it is possible to obtain plate dynamic curves in the form of a group of lines radiating from a single point at different angles. The limiting cases are $R_a = 0$, where the curve runs parallel to the current axis, and $R_a = \infty$ (open plate circuit), where the curve coincides with the voltage axis.

In order to establish a connection between the values of grid voltage and the values of plate current and plate voltage, we plot on a single graph a family of static plate characteristic curves and a load line corresponding to the given load resistance R_a and plate-supply voltage E_a (Fig. 7-3). The points of intersection of the static characteristic curves with the load line determine the voltage on the grid for each value of plate current. Under dynamic operation conditions, to each value of grid voltage there correspond uniquely determined values of plate current and plate voltage. Consequently, in a dynamic regime, the plate current and plate voltage are functions of the grid voltage.

The dependence of plate current upon grid voltage under dynamic conditions is graphically illustrated by the dynamic plate-grid

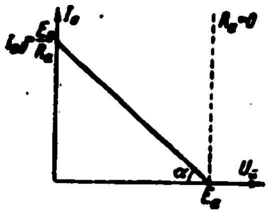

Fig. 7-2. Dynamic plate curve for triode (load line).

(working) characteristic curve. This characteristic (Fig. 7-4) may be plotted on the basis of the points of intersection of the load line and the static plate characteristic curves. If we compare the working

characteristic with the static plate-grid curves plotted on the same graph, it is clear that its slope is less than the slope of the static characteristics. The reason for this is that the increase in plate

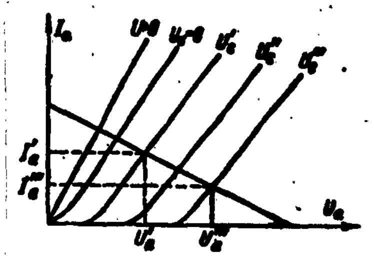

Fig. 7-3. Connection between static plate curves and load line.

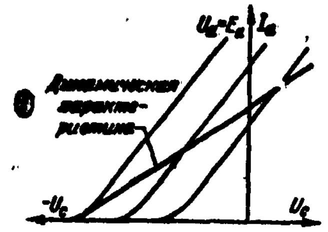

Fig. 7-4. Dynamic plate-grid (working) curve.
1) Dynamic characteristic.

current accompanying an increase in grid voltage under dynamic conditions is less than under static conditions, owing to the simultaneous drop in plate voltage.

The initial point of the dynamic characteristic coincides with the initial point of the static characteristic taken at $U_a = E_a$, since with zero current in the plate circuit, the voltage drop across the load also will be zero and the plate voltage will equal E_a.

7-3. WORKING-CHARACTERISTIC SLOPE.

In utilizing triodes, it is essential to know the slope of the working characteristic, i.e., the quantity determining the change in plate current with a change in grid voltage under dynamic conditions. Let us determine the slope of the characteristic (the transconductance) along the section included between two static characteristics taken at plate voltages U'_a and U''_a (Fig. 7-5). The slope of the working characteristic on this section is defined by the ratio

$$S_d = \frac{I''_a - I'_a}{U''_c - U'_c} = \frac{\Delta I_a}{\Delta U_c}. \qquad (7\text{-}8)$$

If the tube operates under static conditions, a change in grid

voltage to U''_c would cause a change in plate current to the quantity I_{a0}. Let us designate the plate-current increment under static conditions by

$$\delta I_a = I_{a0} - I'_a.$$

In accordance with Eq. (6-17) the static transconductance is

$$S = \frac{\delta I_a}{\Delta U_c}.$$

The segment AB (Fig. 7-5) determines the difference between the increments of current under static and dynamic conditions for precisely the same change in grid voltage:

$$\overline{AB} = \delta I_a - \Delta I_a. \qquad (7\text{-}9)$$

The magnitude of this difference determines the drop in plate voltage from U'_a to U''_a owing to the voltage drop across the load as the plate current rises by an amount ΔI_a. Clearly,

$$U'_a - U''_a = \Delta I_a R_a. \qquad (7\text{-}10)$$

On the other hand, a change in plate voltage with a change in

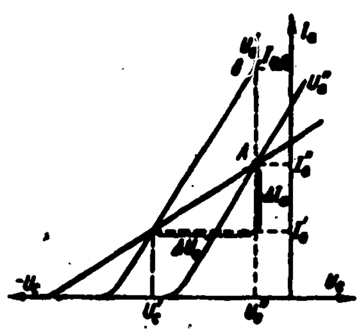

Fig. 7-5. Derivation of the working-characteristic slope.

plate current may be found knowing the plate resistance of the triode:

$$\frac{U'_a - U''_a}{\delta I_a - \Delta I_a} = R_i. \qquad (7\text{-}11)$$

Substituting into this expression the value of the plate-voltage change from Eq. (7-10), and simplifying, we obtain an expression for the quantity ΔI_a:

$$\Delta I_a = \frac{\delta I_a R_i}{R_a + R_i}. \tag{7-12}$$

In order to determine the dynamic (working) transconductance, we substitute (7-12) into (7-8):

$$S_d = \frac{\delta I_a R_i}{\Delta U_c (R_a + R_i)}. \tag{7-13}$$

Since $\delta I_a/\Delta U_c$ = S (static transconductance), we have finally

$$S_d = S\frac{R_i}{R_i + R_a}. \tag{7-14}$$

It is clear from this equation that the working transconductance is the lower the higher the load resistance. This is explained by the fact that a smaller change in plate current is needed with a higher load in order to obtain precisely the same change in plate voltage. The working transconductance is always less than the static transconductance and becomes equal to it where the load resistance is zero. It should be emphasized that the quantity S_d characterizes not the triode, but the amplifier stage (see Sec. 7-4).

7-4. OPERATION OF A TRIODE IN AN AMPLIFIER CIRCUIT.

The simplest amplifier circuit using a triode is shown in Fig. 7-6. The alternating voltage u_c is applied to the grid of the tube (ampli-

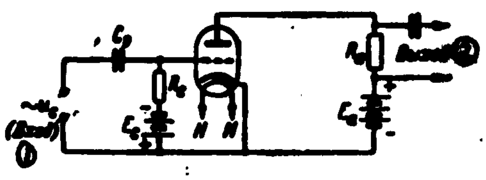

Fig. 7-6. Simplest amplifier circuit using a triode.
1) Input; 2) Output.

fier input) through the isolating capacitor C_r*. At the same time, a constant voltage is applied to the grid from a source E_c; this is called the bias voltage. The isolating capacitor C_r protects the source E_c from being short-circuited across the device supplying alternating voltage to the grid, since this device normally is a transformer winding or a coil, and has a low DC resistance. The load resistance R_a is connected into the plate circuit of the tube, as is the plate-supply source E_a.

At any instant of time, the tube plate current is determined by

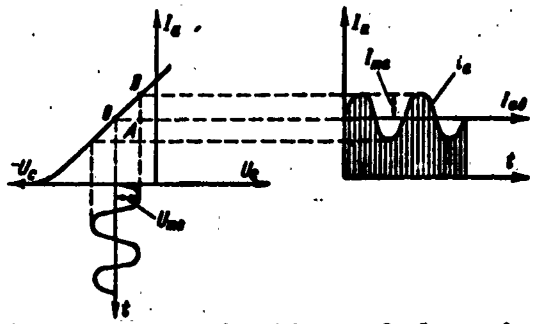

Fig. 7-7. Derivation of dynamic amplification factor.

the sum of the bias voltage and the instantaneous value of the alternating voltage u_c. In this case, the voltage on the grid fluctuates about the bias voltage (Fig. 7-7).

The point O on the dynamic plate-grid curve, corresponding to the bias voltage, is called the operating point.

If the bias voltage alone is applied to the grid, a constant current I_{a0} will flow in the plate circuit. When a signal voltage u_c is applied to the grid together with the bias voltage, an alternating plate-current component i_a will appear in the plate circuit. The plate-current change with time in the presence of both bias voltage and the alternating voltage, called the signal voltage, on the grid is

* [C_p - C_r - $C_{razdelitel'nyy}$ - $C_{isolating}$.]

shown by the hatched section on Fig. 7-7. The current in the plate circuit may be represented as the sum of the DC component I_{a0} and the AC component i_a.

Assuming that along the working section the plate-grid characteristic is linear, we find the ratio of the amplitude of the AC component of the plate current I_{ma} to the signal-voltage amplitude U_{mc}, using the triangle OAB in Fig. 7-7:

$$\frac{I_{ma}}{U_{mc}}=S_d. \tag{7-15}$$

The AC plate-current component sets up a changing voltage drop across the load. The amplitude of the alternating voltage across the load equals the product of the amplitude of the AC current component and the load resistance:

$$U_{mR}=I_{ma}R_a. \tag{7-16}$$

Thus, applying a signal voltage with an amplitude U_{mc} to the input of the circuit – the grid of the tube – we obtain at the output (across the load) an alternating voltage with an amplitude U_{mR}. If $U_{mR} > U_{ms}$, the circuit will operate as a voltage amplifier.

The ratio of the alternating-voltage amplitude at the output to the amplitude of the signal voltage is called the dynamic gain of the amplifier stage

$$\mu_d=\frac{U_{mR}}{U_{mc}}. \tag{7-17}$$

The dynamic gain shows by how much the voltage amplitude at the amplifier output is larger than the signal amplitude at the input.

Substituting for the quantity U_{mR} in Eq. (7-17) its value from (7-16), we obtain:

$$\mu_d=\frac{I_{ma}R_a}{U_{mc}}. \tag{7-18}$$

since

$$\frac{I_{ma}}{U_{mg}} = S_d,$$

then

$$\mu_d = S_d R_a. \tag{7-19}$$

This expression may be transformed by substituting in the value of S_d from (7-14):

$$\mu_d = S \frac{R_i}{R_i + R_a} R_a. \tag{7-20}$$

since $SR_i = \mu$, then

$$\mu_d = \mu \frac{R_a}{R_a + R_i}. \tag{7-21}$$

If the change in signal voltage lies on the linear section of the dynamic plate-grid characteristic curve, then the alternating voltage at the output will be similar to the signal voltage, i.e., amplification will be obtained without distortion. In many triode applications, it is extremely important to obtain undistorted amplification; this is true, for example, in audio amplifier circuits in radio receiving devices, amplification in measuring instruments, etc. Thus, when used in such circuits, the bias voltage on the triode grid is set so that the operating point lies at about the center of the linear section of the curve.

In the circuit described, we have considered the amplification of an alternating voltage. In many cases, it is necessary to obtain across the load a rather high-power alternating current. Examples of such loads are radio-receiver loudspeakers, radio-transmitter antennas, etc. In such cases, it is possible, by applying a low-power input signal, to obtain at the output electrical oscillations of higher power.

In this case, the energy of the DC source is converted into AC energy developed across the load.

In the majority of power amplifiers, circuits are utilized in which the DC component does not appear in the load. Examples of such circuits are circuits with transformer coupling and amplifiers with parallel power supply (Fig. 7-8). In these circuits, the basic power losses in the plate circuit, determined by the DC component, equal the power dissipated on the plate of the tube.

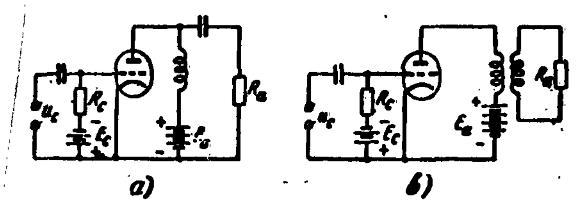

Fig. 7-8. Standard power-amplifier circuits.
a) With parallel power supply;
b) transformer-coupled.

In the circuits discussed, the magnitude of the plate voltage in any instant of time will equal the difference between the plate-supply-source voltage and the instantaneous value of the alternating voltage across the load:

$$u_a = E_a - i_a R_a, \tag{7-22}$$

while the plate current will be the sum of the DC component I_{a0} and the instantaneous value of the AC component i_a:

$$I_a = I_{a0} + i_a. \tag{7-23}$$

since from (7-22)

$$i_a = \frac{E_a - u_a}{R_a}.$$

then

$$I_a = I_{a0} + \frac{E_a - u_a}{R_a} \qquad (7\text{-}24)$$

(for a transformer-coupled amplifier, R_a will be the equivalent resistance of the load, taken from the relationship

$$R_a = n^2 R_n, * \qquad (7\text{-}25)$$

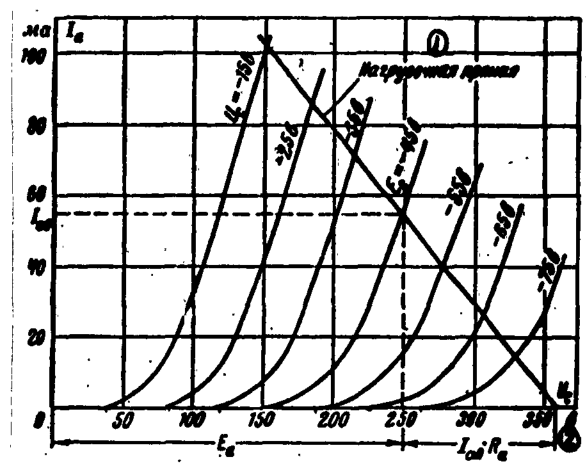

Fig. 7-9. Plot of load line for circuit in which the DC component does not pass through the load. 1) Load line; 2) volts.

where R_n is the actual load resistance, and n is the transformer turns ratio).

Equation (7-24) is the equation of the dynamic plate characteristic of the tube (load line) under given dynamic conditions. The construction of this characteristic is shown in Fig. 7-9. The load line should pass through the point with the coordinates I_{a0} and E_a. This point corresponds to zero instantaneous value of the AC component of the plate current, and lies on the static characteristic taken at a grid voltage equal to the bias voltage E_c. A second point may be found on the line by setting the plate current in (7-24) equal to zero.

---

* [$R_н$ - R_n - $R_{nagruzka}$ - R_{load}.]

Where $I_a = 0$, the plate voltage will equal:

$$U_{a0} = E_a + I_{a0}R_a. \tag{7-26}$$

Plotting the plate-voltage value obtained on the axis of abscissas, we obtain a second point on the curve. The slope of the characteristic curve α may be found from the relationship

$$\operatorname{tg} \alpha = \frac{1}{R_a}. \tag{7-27}$$

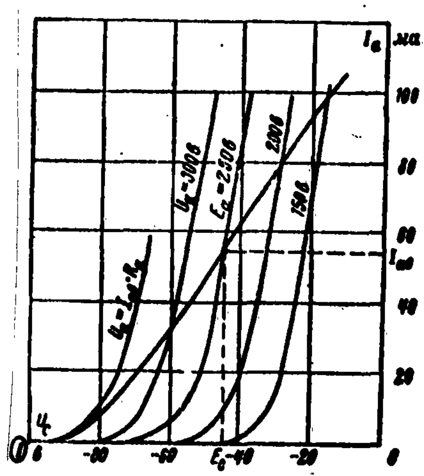

Fig. 7-10. Plot of working characteristic.
1) Volts.

The dynamic plate-grid (working) characteristic curve may be plotted as in the case of a series-connected resistive load in the plate circuit of the tube (Fig. 7-10). The initial point of the working characteristic coincides with the initial point of the static plate-grid characteristic, taken at $U_a = E_a + I_{a0}R_a$. The point of intersection of the dynamic characteristic curve with the static curve, taken at $U_a = E_a$, corresponds to a plate current of I_{a0} and a grid voltage equal to E_c, and is the operating point for the case under consideration.

The dynamic transconductance may be found from the formula already

presented

$$S_d = S \frac{R_i}{R_i + R_a}.$$

It is clear from a consideration of the dynamic characteristics for triodes in amplifier circuits, that for exactly the same value of plate-supply voltage E_a, the range of grid-voltage fluctuations (maximum signal-voltage amplitude) turns out to be larger where the DC plate-current component does not create a voltage drop across a load.

For voltage and power amplification at high frequencies, over a narrow frequency band, so-called tuned amplifiers are used, which have in their plate circuit, as a load, a tuned circuit whose natural resonant frequency equals the signal frequency. At resonance, the circuit presents a pure resistance to the alternating component of the plate current; this is called the resonance impedance, and equals:

$$Z_{pez} = \frac{\omega_0^2 L^2}{R} = \frac{L}{CR}.* \qquad (7\text{-}28)$$

where $\omega_0^2 = 2\pi f_0$ is the angular velocity of the natural circuit oscillations.

The resistance of a circuit to the DC component of the plate current may be neglected. The energy supplied to the tuned circuit from the plate circuit of the tube supports undamped oscillations in the circuit. The equation of the plate current and the load line for such an amplifier are constructed in analogy to the power amplifiers described:

$$I_a = I_{a0} + \frac{E_a - e_a}{Z_{pez}}. \qquad (7\text{-}29)$$

---

* [Z_{pez} - Z_{rez} - $Z_{rezonansnyy}$ - $Z_{resonance}$.]

In an amplifier circuit, only a portion of the power supplied by the source is developed as RF power across the load, and the remaining power is dissipated without performing useful work. For a power-amplification circuit, an important quantity is the plate-circuit efficiency, defined as the ratio of the useful RF power across the load to the power supplied by the source.

In the amplifier circuit considered, the magnitude of the RF power equals:

$$P_{\kappa}=\frac{1}{2}I_{ma}U_{mR}. \text{ *} \tag{7-30}$$

The power delivered by the plate-supply source is:

$$P_{0}=E_{a}I_{a0}. \tag{7-31}$$

From this, the plate-circuit efficiency of the amplifier is

$$\eta=\frac{P_{\kappa}}{P_{0}}=\frac{1}{2}\frac{I_{ma}U_{mR}}{I_{a}E_{a}}. \tag{7-32}$$

Under the conditions obtaining where the current through the tube and the load flows for the entire cycle of alternating signal voltage, the amplitude of the alternating component of the current cannot be greater than the DC component I_{a0}, and the amplitude of the alternating voltage across the load cannot exceed E_a. Consequently, the efficiency of an amplifier operating under such conditions cannot be greater than 0.5. In practice, it ranges from 0.3-0.35.

For a given input-signal amplitude, the value of the RF power is determined from the equality

$$P_{\kappa}=\frac{(\mu U_{mc})^2R_a}{2(R_i+R_a)^2}. \tag{7-33}$$

---

* [P_k - $P_{kolebatel'nyy}$ - P_{RF}]

obtained from (7-30). Differentiating, and setting the derivative equal to zero, it is possible to obtain a relationship between R_1 and R_a, in which the maximum power is developed across the load. The maximum RF power for a given U_{mc} is obtained where $R_a = R_1$, and equals:

$$P_{\kappa.\text{макс}} = \frac{1}{4}\mu S U^2_{mc}. \qquad (7\text{-}34)$$

The power found from expression (7-34) is not the maximum RF power which may be obtained at the output of an amplifier where the values of load resistance and signal-voltage amplitude are simultaneously set, and where distortion is minimal. Analysis of power-amplifier circuits has shown that the maximum RF power with minimum distortion may be obtained where $R_a = 2R_1$.

From the point of view of obtaining maximum efficiency, it is desirable to increase the size of the load resistance; this leads, however, to a sharp drop in the useful RF power.

An increase in efficiency without a drop in RF power may be obtained by increasing the ratio of the amplitude of the alternating-plate-current component to that of the DC component. In amplifiers in which the presence of distortion in the amplified signal is of no great importance, and in tuned amplifiers, where the tuned circuit develops the required harmonic component of the alternating voltage, an increase in the quantity $\gamma = I_{ma}/T_{a0}$ is achieved by changing the operating conditions of the tube. The bias voltage may be so chosen that the plate current flows through the tube only during a portion of the alternating signal-voltage cycle. Such an operating regime of the tube is called operation with plate-current cut-off. Depending upon the position of the operating point and the magnitude of the amplified-signal amplitude in comparison with the plate-current cut-off voltage, three classes of amplification are distinguished. +

If the bias voltage is so chosen that the current flows through the tube during the entire cycle of alternating signal voltage, this amplification regime is called class A operation (Fig. 7-11a). This type of operation is most frequently found in circuits for receiving-equipment amplifiers, where tubes normally operate with no excursion into the positive grid-voltage region and, consequently, with no grid

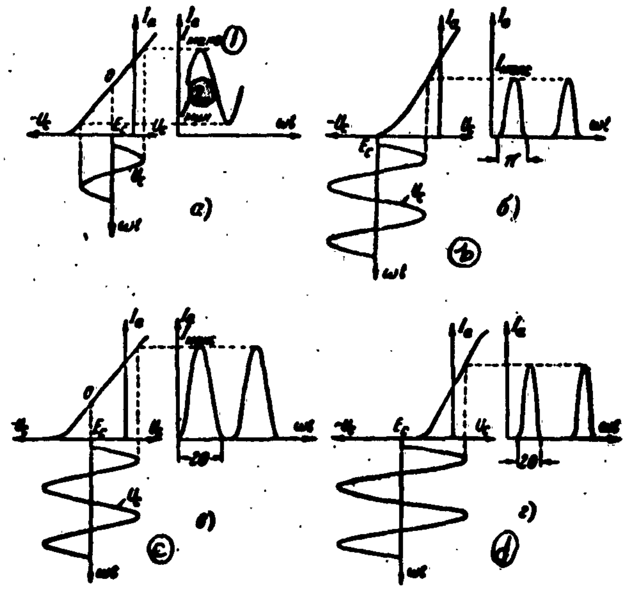

Fig. 7-11. Graphs of triode operation in various classes of amplification.
1) Max; 2) min.

current.

Where a tube is operated with a bias voltage equal to the plate-current cut-off voltage (Fig. 7-11b), this mode of operation is called

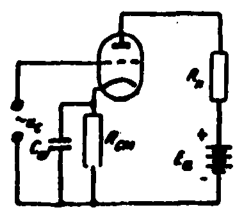

Fig. 7-12. Circuit for amplifier with self-bias.

class B operation. When a tube operates in this regime, the current through it flows only during half of the signal-voltage cycle. During the negative half-cycle, the tube is cut off.

An intermediate situation is represented by class AB operation (Fig. 7-11c), where the tube is cut off by the alternating signal voltage for a period of time that is less than half of a cycle. In this case, a portion of the AC plate-current component is "cut off." The duation of the plate-current pulse is characterized by the cut-off angle θ. Clearly, in class B operation, $\theta = 90^\circ$, while in class AB operation, $180 > \theta > 90^\circ$. In certain circuits, there is an excursion into the region of positive control-grid voltages. In this case, grid current flows; where this current is considerable, noticeable distortion in the shape of the amplified signal appears. This type of operation is indicated by attaching a subscript 2 to the letter designating the class of operation. For example, if the bias voltage equals the plate-current cut-off voltage, while the signal amplitude is greater, this type of operation is called class B_2 operation.

If the operating point of the tube lies to the left of the initial point of the dynamic characteristic curve, this is called class C operation (Fig. 7-11d). When a tube is operated in this class, the length of the plate-current pulse turns out to be less than one-half cycle, i.e., $\theta < 90^\circ$. In amplifiers of radio-receiving equipment, with the exception of telegraph equipment, this type of operation, as a rule, is not used.

In order to establish the required tube operating conditions, the appropriate bias voltage should be applied to the grid. This voltage is supplied either from a special DC source, which is normal for equipment operating with batteries or storage batteries, or else so-called self-biasing is used in which the bias voltage appears across a resistor in

the cathode circuit (Fig. 7-12). The resistor R_{sm}* in the cathode circuit is shunted by a large enough capacitor C_{sh}**. The DC component of the plate current flows through resistor R_{sm}, and the AC component is shunted by the capacitor. The DC component of the cathode current, flowing through the resistor in the cathode circuit, creates a voltage drop across it. Since this resistor is also connected in series with the grid circuit, a voltage will be applied to the grid equal to the voltage drop across the resistance. The size of the resistor in the cathode circuit is normally considerably less than the load resistance and, consequently, it has little effect upon the dynamic characteristic curve of the tube. Clearly, this method of supplying bias voltage may be utilized where there is a large enough DC component of the cathode current. An advantage of self-biasing is the fact that there is some compensation for variations in tube parameters when tubes are changed, owing to the fact that the operating point is shifted to the right when the cathode current decreases, and to the left when it rises.

7-5. OPERATION OF A TRIODE IN AN OSCILLATOR CIRCUIT.

In tuned amplifiers, the DC energy furnished by the plate-supply source is converted into AC energy at a given frequency determined by the parameters of the tank circuit connected into the plate circuit. Consequently, such a circuit may serve as a generator of alternating current. A circuit similar to a resonant-amplifier circuit needs a source of alternating grid voltage in order to operate; this voltage must be at a frequency equal to the natural frequency of the tuned circuit. Such an oscillator is called an independently excited oscillator. A self-excited oscillator may serve as the source of alter-

---

* [R_{cm} - R_{sm} - $R_{smeshcheniye}$ - R_{bias}.]
** [C_{s} - C_{sh} - $C_{shuntiruyushchiy}$ - C_{shunt}.]

nating voltage (Fig. 7-13). In a self-excited oscillator, the alternating voltage on the grid is supplied from the tank circuit itself, as a result of coupling between the grid circuit and the tank circuit. The coupling may be inductive, capacitive, or of the tapped-coil variety

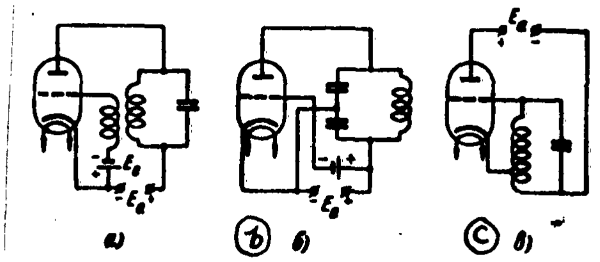

Fig. 7-13. Oscillator circuits, self excited.
a) Inductive coupling; b) capacitive coupling; c) tapped-coil coupling (Hartley oscillator).

(Hartley circuit).

For such an oscillator to operate, the voltage applied to the grid should be opposite in phase to the plate voltage. The plate supply source may be connected either in parallel with the tank circuit or in series with it. Just as in the case of amplifiers, the source of fixed bias in the grid circuit may be replaced by a self-biasing link.

In order to obtain maximum efficiency in oscillators and power amplifiers for radio frequencies, class C operation is normally used, i.e., the tubes are operated with a grid bias whose absolute value exceeds the plate-current cut-off voltage. In this case, the current through the tube flows for a portion of the cycle that is less than a half cycle, corresponding to the positive half cycle of alternating grid voltage. Since the AC component of the plate voltage is opposite in phase to the alternating grid voltage, consequently, the plate current flows at minimum values of plate voltage, which decreases the power dissipated by the tube plate, and increases the plate efficiency of the oscillator (power amplifier). In other words, the ratio of the

AC-component amplitude for the plate current to the value of the DC component entering into the expression (7-32) for the efficiency:

$$\eta = \frac{1}{2}\frac{I_{ma}U_{mk}}{I_{a0}E_a}.$$

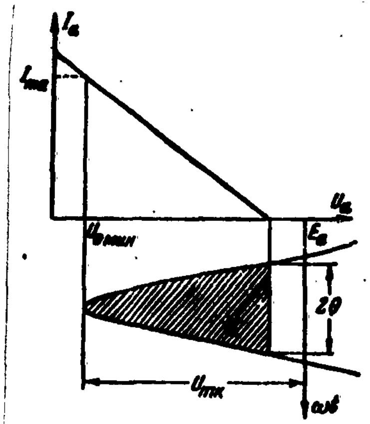

Fig. 7-14. Dynamic plate characteristic curve of oscillater tube operating in class C (cross-hatched section represents a portion of the cycle of alternating voltage in the tuned circuit during which current flows through the tube).

increases in class C operation in comparison with the value for tubes operating without plate-current cut-off.

In addition, in order to increase the efficiency of the plate circuit of an oscillator, it is necessary to try to increase the plate-voltage utilization factor $\xi = U_{mk}/E_a$. The RF power in the tuned circuit and the plate-supply voltage, added together, determine the magnitude of the plate voltage. Since the alternating control-grid voltage of the tube and the voltage in the tank circuit are opposite in phase, at maximum grid potential, the plate voltage will be at a minimum, equal to the difference between the plate-supply voltage and the amplitude of the voltage across the tank circuit:

$$U_{a.min} = E_a - U_{mk}, \qquad (7\text{-}35)$$

as is shown in Fig. 7-14. The closer the plate-voltage utilization factor comes to unity, the lower the value of $U_{a.min}$.

In order to obtain maximum oscillator plate-circuit efficiency, the minimum plate voltage should be reduced to the least permissible value. The utilization of the maximum possible range of variation of plate voltage corresponds to an excursion into the region of positive

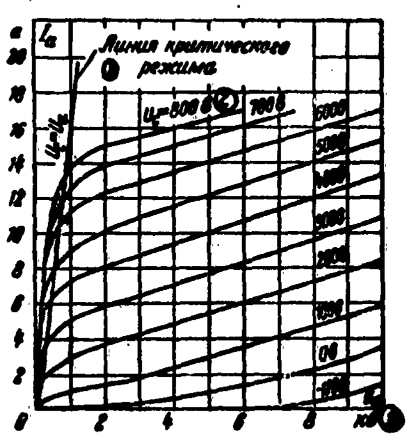

Fig. 7-15. Plate characteristic curves for GU-10A transmitting triode.
1) Critical-operation line;
2) volts; 3) kv.

grid potential. This means that the tube operates with considerable currents in the control-grid circuits, which is a very important feature of transmitting tubes.

Consequently, the basic section of a dynamic plate-grid characteristic curve for transmitter tubes should lie in the region of positive grid voltages, and maximum interest is presented by static characteristics corresponding to this mode of tube operation. The static plate characteristics of triodes, taken at positive values of grid voltage, have two sections: a section with a rapid rise in plate current as plate voltage increases, and a flat section. The nature of the change in plate current is determined by the current distribution

in the tube: the steep section of the characteristic corresponds to electron-return operation, and the flat section to direct-interception operation. At high positive grid potentials, in direction-interception operation, changes in plate voltage have relatively little effect upon the magnitude of the plate current, since the relative change in the effective potential is slight. Typical plate curves for a transmitting triode are given in Fig. 7-15.

Figure 7-16 shows sample curves for a transmitting-tube operating regime. As is clear from the graphs, for a portion of the cycle, plate-current pulses flow through the tube; they may be decomposed into a series, representing the sum of the DC component and a series of sinusoidal harmonics. Clearly, the RF power developed in the tank circuit will depend upon the amplitude of the harmonic component to which the tank circuit is tuned. If the tank circuit is tuned to the frequency of the excitation voltage, there will be power amplification at this frequency, and the magnitude of the RF power is determined by the amplitude of the first harmonic. If the resonant frequency of the tank circuit is twice or three times the excitation-voltage frequency, the RF power in the tank circuit will be developed in accordance with the second or third harmonic of the plate current. In this case, oscillations at double or triple the frequency will be generated, and such circuits are called frequency multipliers.

When a transmitting tube is operated with grid currents, considerable power will be developed at the control grid, at the expense of the source of excitation voltage. In order to decrease this power consumption, it is desirable that the required plate-current amplitude be reached at a lower excitation-voltage amplitude, which corresponds to high dynamic transconductance and, consequently, high static transconductance for the tube. Thus, high transconductance is a most im-

portant quality of transmitting tubes.

A second factor affecting the power developed at grids is the current distribution in the tube. The higher the current-distribution factor for a tube, the lower the grid current corresponding to a given plate-current amplitude, and the less power developed at the grid.

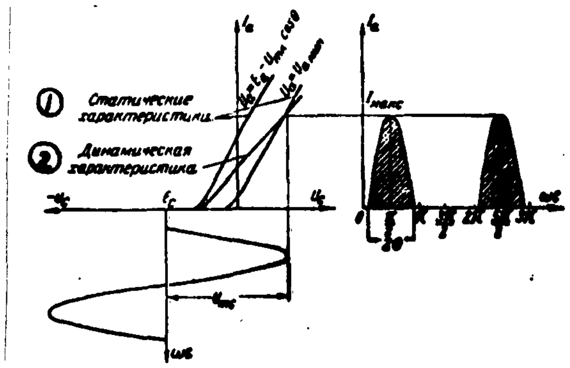

Fig. 7-16. Sample graphs of transmitting tube operating conditions.
1) Static characteristic; 2) Dynamic characteristic.

Under dynamic conditions, for a small portion of a cycle of the alternating voltage, the relationship between the plate voltage and the grid potential may be such that the tube will operate under electron-return conditions. This mode of operation corresponds to a sharp rise in triode grid current with a simultaneous drop in plate current. In this case, there is a considerable increase in grid-circuit power losses. This type of operation of a transmitting tube is called the overvoltage mode. Transmitting tubes are operated under overvoltage conditions in cases where it is especially important to obtain high plate efficiency by increasing the plate-voltage utilization factor which decreases the heat load on the plate. If the plate voltage over an entire cycle remains larger than the voltage corresponding to the transition to electron-return current distribution, the operating mode of the tube is

called the undervoltage mode. The intermediate value of minimum plate voltage corresponds to the critical operating mode for the tube.

In order for a tube to operate in an oscillator circuit with no excursions into the electron-return region, the working section of the dynamic plate characteristic curve for the tube should not intersect the rising sections of the static plate characteristic curves. Thus, as a rule, for the family of static plate characteristics, we construct a graph of the critical operating mode, which for triodes takes the form of a curve showing the plate current as a function of plate voltage where the grid potential is equal to the plate voltage, as shown in Fig. 7-15. The steeper the critical-mode line, the higher the plate-voltage utilization factor may be made without going into the overvoltage mode. For the majority of transmitting tubes, the critical-mode line coincides with the steep portion of the plate characteristics.

The power developed at the grids of transmitting tubes heats them up to a high temperature, which may lead to deformation of the grids, increased evolution of gases from grids, dispersion of material from turns and traverses, and even to overheating of the turns. In addition to the power developed at grids owing to grid current, the grids of transmitting tubes are additionally heated owing to the power radiated from the interior surfaces of plates that are heated to very high temperatures, and the power radiated by cathodes. As a result, the power that may be obtained from a tube in an oscillator circuit is frequently limited not by the permissible plate dissipation, but by the permissible grid dissipation.

7-6. TRIODE TYPES AND APPLICATIONS.

For use in the various types of circuits that have been discussed

in part in the preceding paragraph, triodes should have different characteristics.

As do all electron tubes, triodes fall into two large groups — receiving-amplifying (low-power) triodes and transmitting (high-power) triodes. Receiving-amplifying triodes are classified, as a rule, on the basis of amplification factor, and transmitting triodes according to power and applications.

In low-frequency power amplification circuits, in order to obtain high enough RF power, it is necessary to utilize tubes having, as follows from Eq. (7-34), as large a product $S\mu$ as possible together with a left-shifted characteristic, which makes it possible to utilize a high-amplitude driving voltage. In order to obtain a left-shifted characteristic for tubes having high gains, however, a considerable increase in plate-supply-source voltage is required, which complicates tube and circuit design. In the majority of receiving-amplifying-triode types, the rated plate voltage does not exceed 250-300 v. Thus, low-μ triodes with gains on the order of 4-10 are used for low-frequency power-amplification circuits; they have a left-shifted characteristic and relatively high transconductance. Since the product $S\mu$ for such triodes may not be very high, they are utilized in circuits designed for small amounts of RF power — not more than 3-5 watts.

Such receiving-amplifying triodes have maximum plate dissipations on the order of 10-15 watts. Triodes are considerably more frequently utilized in low-frequency voltage amplification circuits. As follows from expression (7-19), in order to obtain high voltage gain, it is desirable to have a high dynamic transconductance S_d, and a high plate-load resistance. Since the load resistor in such amplifiers normally equals 2-3 R_1, the tubes should have a large plate resistance, which corresponds to a high amplification factor, which in such tubes may

range from 50 to 100. A further increase in gain with no change in plate voltage does not make it possible to obtain the desired values of transconductance, owing to a drop in the effective potential. The plate resistance of high-μ triodes is on the order of several tens of kilohms.

Medium-μ triodes are very common; their amplification factors range from 20 to 40. These triodes may be successfully employed in low-frequency voltage amplification circuits; they have higher transconductance than high-μ triodes and plate resistances of 10-15 kohm. Medium-μ triodes are in universal use, and are used in very different types of circuits — in power amplification circuits for low values of RF power, in pulse generators for special pulse shapes, in computer circuitry, etc.

Transmitting triodes, designed for use in RF power amplification circuits and as high-power RF oscillators, should have a high value of plate dissipation, and large cathode emission. Such tubes are normally operated class C, and operate in the positive grid-voltage region. In order to avoid the necessity for the utilization of high-amplitude driving voltage and a bias voltage that is large in absolute value, transmitting triodes normally have "right-shifted" characteristics, corresponding to a high amplification factor (on the order of 40-80).

Transmitting triodes operating in low-frequency power-amplification circuits (modulator triodes) basically are subject to the same requirements as are receiving-amplifier triodes. High-power modulator triodes normally operate in class A, without going into the positive grid-voltage region. The name "modulator" triode is somewhat inaccurate, since modulation proper takes plate in tubes operating in intermediate stages of RF power amplification, while modulator tubes are utilized to increase the power of the modulating signal.

Figure 7-17 shows the simplest plate-modulation circuit. In this circuit, the drive voltage is applied to the input of the transmitting tube (RF power amplifier), while alternating audio voltage from the microphone is applied to the modulator input. The modulator tube is used for power amplification of the audio-frequency signal; in this case, the transmitting tube acts as the load for the audio-frequency amplifier and, consequently, the magnitude of this load is determined by the internal resistance of the transmitting tube. As a result, the modulated RF is taken from the output circuit of the transmitting tube. Both tubes use a common plate-supply source, which is protected against audio voltages by an audio choke. The entire modulator section of the

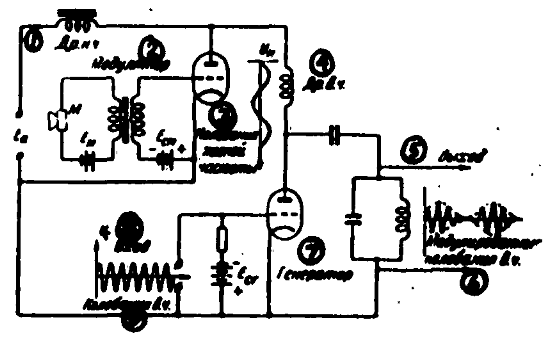

Fig. 7-17. Plate-modulation circuit.
1) Audio choke; 2) modulator; 3) audio signal; 4) RF choke; 5) output; 6) modulated RF; 7) RF tube; 8) input; 9) RF.

circuit is protected against RF voltages by a RF choke in the plate circuit of the transmitting tube.

Thus, the modulator tube operates in a low-frequency power-amplification circuit (in the given case the signal is formed by audio oscillations arriving from the microphone).

Circuits utilizing triodes are greatly affected by tube internal capacitances, which have an especially strong effect at radio frequencies.

The interelectrode capacitances shunt the input and output circuits, resulting in a drop in the dynamic gain of the amplifier circuits, and distortion of transmitting-tube frequency response.

The grid-plate capacitance has an especially strong effect upon circuit operation (the transfer capacitance). It is known from the theory of vacuum-tube amplifiers that the dynamic input capacitance of a tube, which determines the output conductance, is determined by the total input capacitance and the plate-grid capacitance, multiplied by a factor of $\mu_d + 1$:

$$C_{вхд} = C_{c.к} + (\mu_d + 1) C_{a.c}. \qquad (7\text{-}36)$$

This over-all capacitance shunts the plate load of the tube in the previous stage of amplification, which leads to a considerable loss of gain. This phenomenon takes place with greater intensity the higher the frequency of the alternating voltage.

The plate circuit and grid circuit are coupled through the capacitance $C_{a.c}$. In amplifiers having tank circuits, such "feedback" may lead to the appearance of oscillation, which cannot be permitted, since the circuit begins to operate as a self-excited oscillator. The phenomenon of self-excitation is also possible in the absence of tuned circuits in the plate circuit. In such circuits, a tuned circuit may be created by the plate-cathode capacitance, the capacitance between circuit wiring elements, and the inductances of circuit wiring and tube leads.

In standard designs for receiving-amplifying triodes, the interelectrode capacitances may be on the order of 1 micromicrofarad, and in transmitting tubes on the order of tens of micromicrofarads.

A substantial decrease in interelectrode capacitance of triodes as a result of structural changes is very difficult. In the best RF

triodes of standard construction, the transfer capacitance does not decrease below 1 micromicrofarad. Special triode designs, intended for amplification at microwave frequencies, are discussed in Chapter Twelve.

Thus, the basic fundamental disadvantages of triodes are high transfer capacitance, and the impossibility of gaining any considerable increase in plate resistance with high enough transconductance without a noticeable rise in plate voltage. The second drawback is peculiar basically to receiving-amplifying triodes. In transmitting tubes, it is possible to apply high plate voltages such as are needed for obtaining large RF powers.

These drawbacks are to a considerable degree eliminated in more complicated types of tubes — tubes with screen grids.

Chapter Eight

SCREEN-GRID TUBES — TETRODES AND PENTODES

8-1. TETRODE ARRANGEMENT AND OPERATING PRINCIPLES.

The simplest screen-grid electron tube is the four-electrode tube — the tetrode (Fig. 8-1). In this tube, a second grid, called the screen grid, is located in the space between the control grid and the plate. It has a positive potential ranging from 15 to 100 per cent of the nominal plate voltage, depending upon the design and function of the tetrode. The most important advantage of the tetrode lies in the fact

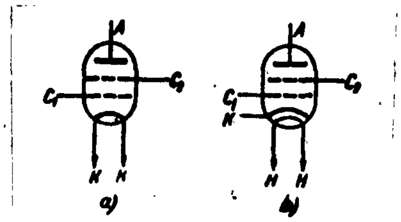

Fig. 8-1. Designations for tetrode elements.
a) With filament; b) with indirectly heated cathode.

that the second grid shields the plate from the control grid, resulting in a considerable decrease in the transfer capacitance of the tube. A comparison of similar designs for triodes and tetrodes shows that the presence of a screen grid decreases the transfer capacitance by factors of tens and hundreds.

The presence of an additional second grid near the first grid increases the effective potential without raising the plate voltage. On the other hand, this second grid weakens the action of the plate electric field on the space charge near the cathode and, consequently, decreases the effect of a change in plate voltage on the cathode current, which increases tube gain. Thus, the introduction of the second grid

makes it possible to raise tube gain in the presence of a quite high effective potential without raising plate voltage.

The first grid of a tetrode is still the control grid. The electron current coming from the cathode, passing through the first grid, is distributed between the plate and the second grid. The screen-grid current is of no use, and thus, in designing tetrodes, it is attempted to minimize it.

8-2. EFFECTIVE POTENTIAL OF A TETRODE.

The electric field in a tetrode is set up by three electrodes — the first and second grids and the plate. Consequently, the magnitude of the currents for the electrodes of a tetrode is a function of the potentials of these electrodes.

The magnitude of the total cathode current is determined by the resultant potential set up in the space between the first grid and the cathode. Just as with a triode, the potential in the plane of the first grid of a tetrode is not the same at different points. Thus, in order

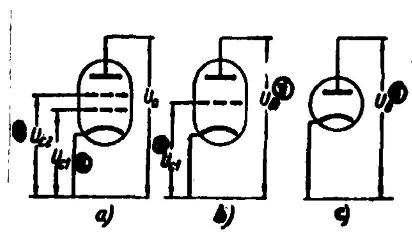

Fig. 8-2. Reduction of a tetrode to an equivalent diode.
a) Tetrodes; b) equivalent triode; c) equivalent diode.
1) U_{c2}; 2) U_{c1}; 3) U_{d2}; 4) U_d.

to compute the magnitude of the cathode current, it is necessary to replace the nonuniform potential in the plane of the first grid with the resultant potential set up at the cathode by a field equivalent to the field set up as a result of the resultant effect of the poten-

tials of the first and second grids and the plate. In order to determine this effective potential, it is possible to utilize a method of successive reduction to the equivalent diode, as was done in the discussion on triodes.

The resultant effect of the plate and the screen grid may be replaced by the effect of an imaginary plate located in the plane of the screen grid (Fig. 8-2), having a potential U_{d2} determined from the well-known formula

$$U_{д2}=U_{c2}+D_2U_a.^{*} \qquad (8\text{-}1)$$

In this formula, D_2 is the penetrance of the second grid, determining the electrostatic effect of the plate on the second grid.

As a result of replacing the plate and screen grid with an equivalent plate, we obtain a triode with a plate voltage U_{d2} and a grid voltage of U_{c1}. Such a triode may be reduced to an equivalent diode with an effective potential:

$$U_д=U_{c1}+D_1U_{д2}, \qquad (8\text{-}2)$$

where D_1 is the penetrance of the control grid, determining the electrostatic effect of the equivalent plate on the cathode.

Here it is assumed that the penetrance D_1 is quite small and, consequently, it may be neglected in the denominator of the effective-potential equation.

Substituting the value of the potential U_{d2} into (8-2), we obtain an equation defining the effective potential of a tetrode:

$$U_д=U_{c1}+D_1U_{c2}+D_1D_2U_a. \qquad (8\text{-}3)$$

---

* [$U_д$ - U_d - $U_{deystvuyushchiy}$ - $U_{effective}$.]

The product $D_1D_2 = D$ is called the total penetrance of a tetrode. The penetrance of a tetrode characterizes the electrostatic effect of the plate of the tetrode on the cathode. Since the quantities D_1 and D_2 are always less than 1, D is much less than 1. The equation for the effective potential of a tetrode finally takes the following form:

$$U_д = U_{c1} + D_1 U_{c2} + DU_a. \quad (8\text{-}4)$$

Since the quantity D is very small, the effect of plate voltage on the space charge near the cathode of a tetrode is slight. Thus, the value of the effective potential is determined basically by the potentials of the grids, and an abbreviated version of the effective-potential equation may be utilized with accuracy sufficient for the majority of tetrodes:

$$U_д \approx U_{c1} + D_1 U_{c2}. \quad (8\text{-}5)$$

Substituting into the three-halves law the tetrode effective potential correct for the equivalent diode, we may find the magnitude of the cathode current.

$$I_к = G U_д^{\frac{3}{2}} \approx G(U_{c1} + D_1 U_{c2})^{\frac{3}{2}}. \quad (8\text{-}6)$$

The geometric coefficient G in the equation is determined from the dimensions of the control grid and the cathode of the tetrode.

8-3. DISTRIBUTION OF CURRENTS IN A TETRODE. PLATE-CURRENT EQUATION.

In the general case, the electron current from the cathode of a tetrode is distributed between the grids and the plate. The relationship between the currents in these electrodes depends upon their potentials and upon geometry. In the particular case of the negative control grid, the current from the cathode is distributed between the screen and the plate. In this case, the control grid may be con-

sidered to be the cathode, delivering the total tetrode current.

The distribution of current between plate and screen is similar to the distribution of current in a triode where the grid is positive. Where the plate voltage is greater than the screen voltage, the current distribution corresponds to direct interception conditions, as discussed in Sec. 6-5. In this case, the current-distribution factor will equal:

$$\frac{I_a}{I_{c2}} = k = C_1 \sqrt{\frac{U_a}{U_{c2}}}. \qquad (8\text{-}7)$$

For plane electrodes, the geometric factor C_1 takes the following form:

$$C_1 = \frac{1}{\alpha}\left(\frac{r_{c2}}{r_a}\right)^{\frac{2}{3}}; \qquad (8\text{-}8)$$

for cylindrical electrodes

$$C_1 = \frac{1}{\alpha}\left(\frac{r_{c2}}{r_a}\right)^{\frac{1}{3}}. \qquad (8\text{-}9)$$

In these formulas, α is the grid fill factor defined as the ratio of the portion of the grid surface occupied by the turns and the traverses to the entire surface. The quantities r_{c2} and r_a are the screen grid-cathode and plate-cathode separations in a flat arrangement, or the radii of the screen grid and plate in a cylindrical system.

The current-distribution coefficient under electron-return conditions, for a tetrode, may be found as in the case of a triode (6-39):

$$\frac{I_a}{I_{c2}} = k = 1{,}4C_1\left(\frac{U_a}{U_c}\right)^2. \qquad (8\text{-}10)$$

Comparing Eq. (8-7) for the current-distribution factor in direct interception with Eq. (8-10), we find the current-distribution factor under electron-return conditions; clearly, as the plate voltage rises, the current-distribution factor and, consequently, the plate current

rise considerably faster under electron-return conditions than under direct interception conditions. This conclusion is illustrated graphically in Fig. 8-3, which shows the current-distribution factor as a function of the ratio U_a/U_{c2} with a constant voltage on the screen grid.

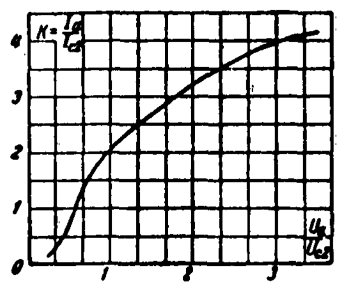

Fig. 8-3. Example of the dependence of the current-distribution factor of a tetrode upon the ratio of the plate voltage to the screen voltage.

The distribution of current in a tetrode may change substantially as a result of secondary electron emission from the screen and plate of the tube. Where the plate and screen potentials are high enough, the magnitude of the secondary-emission current may reach considerable values, leading to distortion of the tube characteristic. Secondary electrons move from an electrode at lower potential to an electrode at higher potential. In such case, the total electron current of an electrode at a lower potential drops, and the current of the electrode at higher potential rises. Thus, functional dependencies of current division may be investigated only with small positive electrode potentials, where the energy of the primary electrons is insufficient to cause noticeable secondary emission.

Where the plate voltage is less than the screen voltage, but high enough for secondary emission to appear, secondary electrons from the plate move to the screen under the action of the electric

field, decreasing the plate current and raising the screen current. The drop in plate voltage under these conditions leads to a rise in the velocity of the electrons arriving at the plate, which in turn causes an increase in the secondary-emission coefficient. As a result, an increase in plate voltage is accompanied by a drop in plate current. A similar phenomenon may occur at the grid under direct-interception conditions.

In order to determine the magnitude of the plate current of a tetrode under given conditions where the control grid is negative, it is necessary to know the current distribution between the screen and plate. Thus,

$$I_a + I_{c2} = I_k$$

and

$$\frac{I_a}{I_{c2}} = k,$$

then

$$I_a = \frac{k}{k+1} I_k. \qquad (8\text{-}11)$$

The cathode current of a tetrode is found from Eq. (8-6); then

$$I_a = \frac{k}{k+1} G U_A^{\frac{3}{2}}. \qquad (8\text{-}12)$$

8-4. STATIC PLATE CHARACTERISTICS OF TETRODES.

The magnitudes of the plate current and grid currents of a tetrode depend upon the potentials of three electrodes: the plate, and the two grids. The characteristics indicating the dependence of these currents upon one of these voltages with the other two held constant are called the static characteristics of the tetrode.

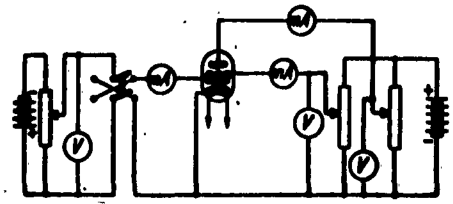

Fig. 8-4. Circuit for studying a tetrode.

In practice, these characteristics may be found with the aid of the circuit shown in Fig. 8-4. In order to simplify the drawing, the heater supply circuit has not been shown. The plate voltage and screen voltage are taken from potentiometers connected in parallel with a common DC supply. The control grid voltage is taken from a potentiometer through a changeover switch that makes it possible to change grid polarity. Milliammeters are used to measure the plate and grid currents. The voltages on the electrodes are found with voltmeters connected directly to the potentiometers.

The curve showing plate current as a function of plate voltage with the grid and screen voltages held constant is called the static plate characteristic of the tetrode. The characteristics showing the grid currents as functions of plate voltage under the same conditions are called the grid-plate characteristics.

Sample plate and grid-plate (screen grid) characteristics of a tetrode are shown in Fig. 8-5. As is clear from the graph, the characteristics have three sections: section I corresponding to the rapid rise in plate current as the plate voltage rises from zero, section II on which we observe a dip in the plate-current characteristic as the plate voltage rises, and section III, along which an increase in plate voltage causes a slight rise in plate current.

The rapid rise in plate current on the first section results from

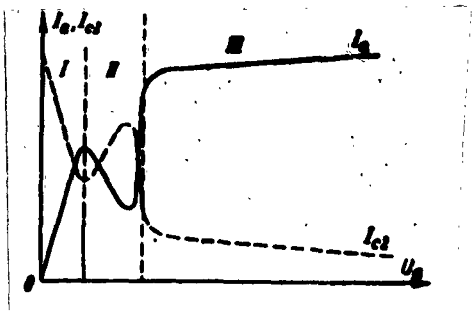

Fig. 8-5. Sample plate and grid-plate characteristics of a tetrode.

the rapid rise in the current-distribution factor as the plate voltage increases under electron-return conditions. On this section, there is no secondary emission from the plate, since the energy of the electrons arriving at the plate, where the plate voltage is small, is inadequate for the appearance of any noticeable secondary emission. When the electrons arriving at the plate obtain a sufficiently high energy, secondary emission from the plate appears; it increases as the plate voltage rises. As long as the screen voltage remains higher than the plate voltage, secondary electrons will pass over to the screen under the influence of the electric field in the plate-screen space; this results in a rise in the screen current and a decrease in plate current. In this case, a rise in plate voltage is accompanied by a drop in plate current: there is a dip in the plate characteristic. This phenomenon is called the dynatron effect.

With a further rise in plate voltage, when it comes near to the value of the voltage on the screen, the dynatron effect ceases to be felt, since all of the secondary electrons return back to the plate and, consequently, have no effect upon the current distribution. When the plate voltage rises above the screen voltage, a direct-interception mode of operation sets in, and the plate current rises slightly with

plate voltage. Thus, the plate characteristic along the third section is sloping.

Since the change in plate voltage has very little effect upon the value of the effective potential of a tetrode, there is almost no increase in cathode current accompanying a rise in plate voltage. Thus, the sum of the plate and screen currents remains almost constant.

The existence of the dynatron effect strongly limits tetrode applications. Under dynamic operating conditions, the plate voltage may become less than the screen voltage, and the tube will operate in the dynatron-effect section. When this happens, undesirable phenomena appear: excitation of spurious oscillations in plate circuits, distortion of the signal being amplified, etc.

8-5. METHODS FOR ELIMINATING THE DYNATRON EFFECT. BEAM TETRODES AND PENTODES.

In order to eliminate the dynatron effect, it is necessary to prevent the movement of secondary electrons from the plate to the screen grid. In order to do this, an electrical field that will return the secondary electrons to the plate should be set up in the space between the plate and the screen.

If a minimum potential is set up in the plate-screen space, primary electrons having high velocities will overcome it and escape the plate. Secondary electrons, having low initial velocities, cannot overcome this minimum potential, and will return to the plate.

It is possible to set up a minimum potential in the plate-screen space, either by increasing the space-charge density in this space, or by introducing an additional electrode into the space that has a potential considerably below the plate and screen potential.

The simplest method for setting up a minimum potential in the screen-plate space when $U_{c2} > U_a$ is to move the plate away from the

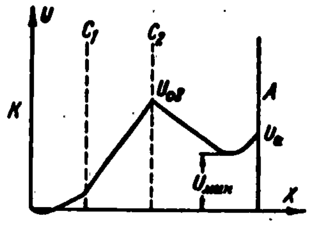

Fig. 8-6. Distribution of potential in a tetrode where $U_a < U_{c2}$.

screen. To clarify this, let us consider the change in potential as we move from the cathode to the plate (Fig. 8-6) in a tetrode (we consider the grid potentials to be their effective potentials). In the screen-plate space, a minimum potential is formed as a result of the space charge set up by the electrons passing through the screen grid. The greater the number of electrons that reach this space during each instant of time, the deeper will be the potential minimum. As the plate is moved away from the screen, the number of electrons in the screen-plate space will rise. In addition, there will be a drop in the resultant potential of the screen grid, and the velocity of the electrons passing through the screen drops somewhat, which also amounts to an increase in electron concentration. All of this leads to a drop in the potential minimum, and places it further away from the plate. Where the current passing through the screen has sufficiently high density, the potential minimum that forms proves to be adequate to cause the slow secondary electrons to return to the plate and, consequently, to suppress the dynatron effect. The required relationship between the screen-plate and cathode-plate separations that will suppress the dynatron effect depends upon the potential of the screen and the electron-current density. In most cases, the condition

$$\frac{r_a - r_{c2}}{r_{c2}} > 10$$

should be satisfied. At the present time, the method described for suppressing the dynatron effect is utilized relatively infrequently, and only in tubes having very small cathode-screen grid spacings (for example, in the 6E5P tetrode). Where the situation is otherwise, removal of the plate to a greater distance leads to a considerable rise in tube dimensions.

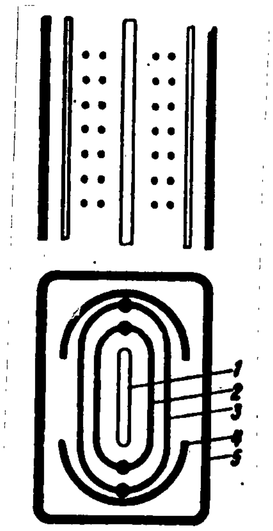

Fig. 8-7. Diagramatic representation of a beam tetrode.
1) Cathode; 2) control grid; 3) screen grid; 4) beam-forming plate; 5) plate.

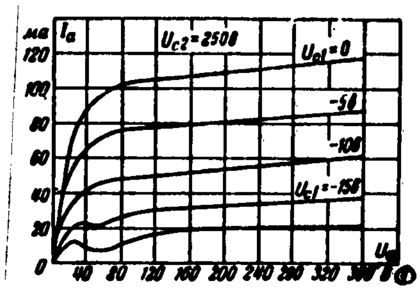

Fig. 8-8. Family of plate characteristics of the 6P6S [6V6GT] beam-power tube.
1) Volts.

A more widely utilized design is the so-called <u>beam-power tetrode</u>, in which a minimum potential is set up in the screen-plate space by increasing the space charge, without moving the plate further away. Beam tetrodes have the following structural peculiarities. Both tetrode grids are made with the same pitch and are so arranged in the tube that the projections of their turns on the cathode coincide, i.e., so that the turns of the screen grid are so located with respect to the cathode

that it is as if they were in the shadow of the turns of the first grid. In the space between the screen grid and the plate, there is a beam-forming plate, having a longitudinal slot (Fig. 8-7). The beam-forming plate is connected to the cathode and has, consequently, zero potential.

As a result of the fact that the turns of both grids are in the same position, the electron stream passing through the control grid suffers almost no scattering along the axis of the tube. Passing through the screen grid, the stream of electrons is laterally compressed by the field of the beam-shaping plates. Acted upon by the field formed by the grid turns, and the field of the slot in the beam-shaping plate, the stream of electrons is formed into electron beams of high density. The high density of the electron charge sets up in the screen-plate space a minimum potential that prevents the movement of secondary electrons from the plate to the screen grid.

Sufficiently high electron density in a beam tetrode may be achieved where the total magnitude of the current in the tube is sufficiently great. At low current values for the cathode, the density of the electrons in the screen-plate space proves to be insufficient for the formation of a potential minimum and the prevention of the dynatron effect. This is clear from the family of plate characteristics for the 6P6S beam tetrode (Fig. 8-8), which takes the form of a group of characteristics taken at constant voltage on the screen grid for several values of control-grid voltage. The characteristics, corresponding to adequately high values of grid voltage (from -10 v to 0), do not have a dynatron dip, which is explained by the high current of the tube — more than 40 ma. At small operating currents (less than 25 ma), corresponding to a grid voltage below -15 v, the characteristics display a well-defined dip. Consequently, the beam-tetrode

construction is designed for operation at relatively large plate currents, and is normally utilized in tubes for power amplification.

One of the considerable advantages of beam tetrodes, increasing their operating economy, is the low screen-grid current with respect to the plate current, which in the majority of beam tubes does not exceed 7-10 per cent of the plate-current value. Such a ratio of plate and screen currents is provided by the matched arrangement of grid turns, which results in a considerable drop in the interception of electrons by the screen grid.

Since with low currents the beam-power construction does not provide suppression of the dynatron effect, tubes are used in such cases that have a third antidynatron (or suppressor) grid located in the space between the plate and the screen grid. Such electron tubes having three grids are called pentodes (Fig. 8-9). The third grid of the pentode is either internally connected with the cathode, and thus has zero potential, or has a separate lead to which may be applied a potential considerably less than the potentials of the plate and screen. The presence of a grid with a low potential between the plate and screen sets up a potential minimum in the space and, consequently, protects the tube against the occurrence of a dynatron effect.

In order not to produce a considerable deterioration in the distribution of currents (increasing the screen current), the third grid is normally made coarse, with large penetrance. In this case, the majority of primary electrons passing through the screen grid arrive at the plate, with the exception of those electrons whose paths are directed directly against the turns of the suppressor grid. Such electrons may be repelled by the suppressor and returned to the screen.

The plate characteristics of pentodes do not have dips caused by the dynatron effect. The feature distinguishing the plate characteristics

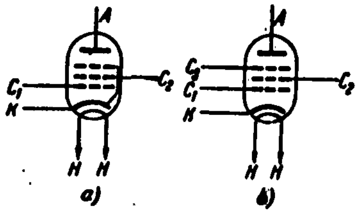

Fig. 8-9. Pentode designation.
a) With a jumper between suppressor and cathode;
b) with separate lead for the suppressor.

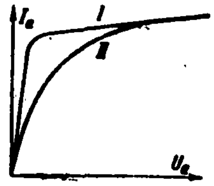

Fig. 8-10. Examples of plate characteristics for beam-power tetrodes (I) and pentodes (II).

of pentodes and beam-power tetrodes consists of the fact that in pentodes, the transition from the steep to the flat segment of the curve curves more smoothly (Fig. 8-10), which limits the range of variation of plate voltage in dynamic operation (decreases the possible plate-voltage use factor). This circumstance is explained by the additional scattering of electrons in the interturn space of the suppressor. The strongly deflected electrons increase the density of the space charge near the plate, slowing down, for example, the transition from electron-return conditions to direct-interception conditions. Scattering of electrons by the suppressor is intensified where they are much scattered in their passage through the screen.

An increase in the slope of the rising portion of the plate characteristic of a pentode may be achieved by increasing the suppressor-grid spacing and by placing it near the plate. The resulting changes in electron trajectories are shown in Fig. 8-11. The increase in spacing and the amount by which the suppressor may be moved nearer to the plate is limited by the possibility of the appearance of a dynatron effect. A beneficial effect results from the utilization of a beam-power system of arranging the control and screen grids; this decreases

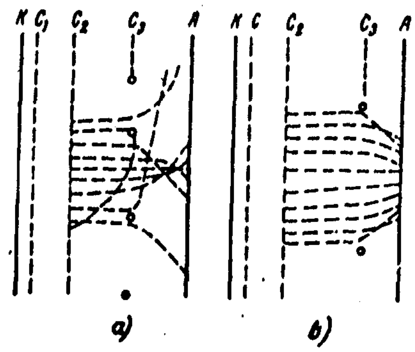

Fig. 8-11. Effect of suppressor-grid geometry on the nature of electron trajectories.

scattering of electrons passing through the screen.

The introduction of a suppressor grid still further decreases the effect of the plate voltage on the cathode space charge. Just as for a tetrode, it is possible to derive an equation for a pentode by successive reduction of the pentode to the equivalent diode:

$$U_A \approx U_{c1} + D_1 U_{c2} + D_1 D_2 U_{c3} + D_1 D_2 D_3 U_a, \tag{8-13}$$

where D_3 is the penetrance of the suppressor grid.

The product of the penetrances of all three grids $D_1 D_2 D_3 = D$ is the total penetrance for the pentode. Although the penetrance of suppressor grids is normally large in comparison with that of the control and screen grids, it has a considerable effect upon the total penetrance of the tube.

8-6. PLATE-GRID CHARACTERISTICS OF TETRODES AND PENTODES.

The static plate-grid characteristics of tetrodes and pentodes are considered to be the curves expressing the plate and grid currents as functions of the control-grid voltage with the plate, screen, and suppressor voltages held constant. In practice, plate-grid character-

istics are also utilized that are indicative of the relationship between plate current and the voltage on the screen with constant voltages on the remaining electrodes. Figure 8-12 shows the family of characteristics $I_a = f(U_{c1})$ and $I_{c2} = f(U_{c1})$, taken for various plate voltages.

The curves for plate current and screen current begin at a single point, corresponding to a cutoff voltage on the control grid. Since the plate voltage has little effect upon the value of the effective potential, the magnitude of the cutoff voltage is basically determined by the screen voltage. As the control-grid voltage rises, the cathode

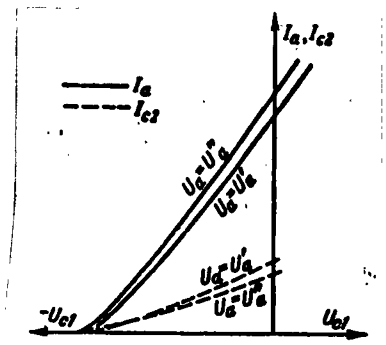

Fig. 8-12. Sample plate-grid and grid (with respect to screen) characteristics of a pentode.

current increases and, consequently, the plate and screen currents increase. Since the plate and screen potentials remain constant, the current distribution does not change either, i.e., the ratio of the plate current to the screen current remains constant.

If the characteristics are taken at some other plate voltage, there will be almost no shift in the starting point of the curves. In pentodes, the change in plate voltage has less effect upon the starting point of the curve than in tetrodes, since the penetrance is lower in pentodes. A change in plate voltage, however, leads to a change in current distribution, and the plate-current characteristic will be

higher with a high plate voltage and the screen-grid current lower than at a low plate voltage.

The family of tetrode and pentode plate-grid characteristics takes the form of a set of curves fanning out from a single point. Pentode characteristics lie within a narrower region than those of tetrodes; this is explained by the fact that the current-distribution coefficient for pentodes changes less as plate voltage rises than is the case with tetrodes. The distribution of current in a pentode may be calculated according to formula (8-7), replacing the suppressor grid with an equivalent plate located in the position of the suppressor and having a potential

$$U_{д3} = \frac{U_{c3} + D_3 U_a + D_{R3} U_{c2}}{1 + D_3 + D_{R3}},$$

where D_3 is the penetrance of the suppressor grid, D_{R3} is the so-called "inverse" penetrance of the suppressor grid — a quantity characterizing the electrostatic effect of the screen grid on the suppressor-plate space.

The current-distribution coefficient in a pentode, under direct-interception conditions, may be determined from the formula

$$\frac{I_a}{I_{c2}} = k = C_1 \left(\frac{U_{д3}}{U_{c2}} \right)^{\frac{1}{2}}. \qquad (8\text{-}14)$$

In this case, it is assumed that the potential of the suppressor grid is either negative or equal to zero and, consequently, there is no current in the suppressor circuit. When the plate current changes by a magnitude ΔU_a, the effective potential in the plane of the suppressor grid changes by a small amount approximately equaling $D_3 \Delta U_a$. Thus, the current-distribution coefficient decreases less in a pentode than in a tetrode for the same given change in plate voltage.

We should note that the presence of a suppressor grid in a pentode

leads not only to suppression of the dynatron effect, but also to the prevention of motion of secondary electrons from the screen to the plate, which increases the plate current of a tetrode where $U_a > U_{c2}$. As we have already stated, secondary-electron emission from the screen grid may prove to have a considerable influence upon the current distribution in a tetrode. In a pentode, in view of the presence of a potential minimum created by the suppressor, the effective secondary emission upon current distribution is practically eliminated.

It should be noted that for the same reasons, the plate characteristic of a pentode is flatter along the section corresponding to direct-interception conditions than that of a tetrode.

Still narrower is the fan produced by the plate-grid characteristics of beam tetrodes for which, under direct-interception conditions, a change in plate voltage has almost no effect upon the current distribution. The position of the plate-grid characteristics is determined basically by the magnitude of the screen-grid voltage. The plate-current cutoff voltage (blocking voltage) may be computed from the condition that the effective potential at the point of origin of the characteristics is zero:

$$U_{10} \approx U_{c10} + D_1 U_{c2} = 0, \tag{8-15}$$

from which

$$U_{c10} \approx -D_1 U_{c2}. \tag{8-16}$$

Consequently, an increase in screen voltage shifts the plate-grid characteristic to the left. The magnitude of the shift of the characteristic is of importance in determining the usefulness of a tube in one or another amplifier circuit. Thus, families of plate-grid characteristics for various screen voltages are more common (Fig. 8-13)t

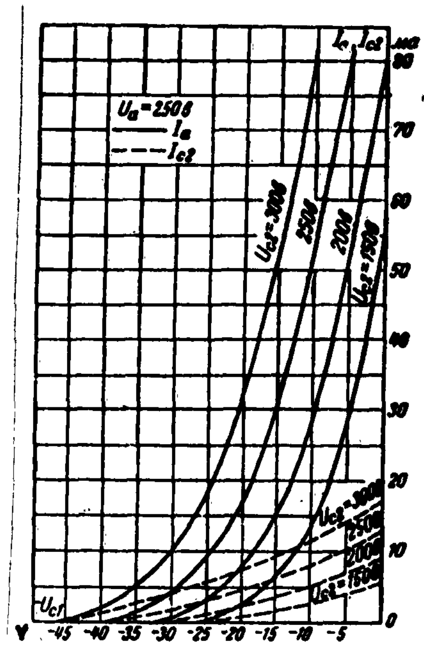

Fig. 8-13. Family of plate-grid and grid (with respect to screen) characteristics of a 6P6S tetrode, taken for various screen voltages.

8-7. STATIC CHARACTERISTICS OF TETRODES AND PENTODES.

The static characteristics of tetrodes and pentodes determining the connection between changes in the basic quantities characterizing tube operation, are, as in the case of a triode, the mutual conductance, the internal resistance, and the amplification factor.

In contrast to the triode, the characteristics of tetrodes and pentodes are determined not only by the degree to which changes in plate voltage and control-grid voltage affect cathode current, but also by the effect of these changes on current distribution. All three characteristics are measured with the screen voltage held constant.

The mutual conductance is defined as the ratio of a change in

plate current to the change in control-grid voltage causing it with the plate and screen voltages held constant. A change in control-grid voltage by ΔU_{c1} causes a change in cathode current by an amount ΔI_k. It is clear that this current increment is distributed between the plate and the screen in accordance with the current-distribution coefficient under the given operating conditions; the increment in screen current and the increment in plate current will add up to the increment in cathode current:

$$\Delta I_a + \Delta I_{c2} = \Delta I_k. \tag{8-17}$$

The ratio of the plate-current increment to the screen-current increment equals the current-distribution coefficient k and, consequently, the plate-current increment may be expressed in terms of ΔI_k:

$$\Delta I_a = \frac{k}{k+1} \Delta I_k. \tag{8-18}$$

The mutual conductance, by definition, equals:

$$S = \frac{\Delta I_a}{\Delta U_{c1}} = \frac{k}{k+1} \frac{\Delta I_k}{\Delta U_{c1}} = \frac{k}{k+1} S_k. \tag{8-19}$$

The quantity $S_k = \Delta I_k / \Delta U_{c1}$ is the mutual conductance of the equivalent triode whose plate is located where the screen is located. The mutual conductance of such a triode is determined in accordance with Eq. (6-19a):

$$S_k = 3{,}5 \cdot 10^{-6} \frac{F_{c1\,\text{эф}}}{r_{c1}^2 \beta^2} U_A^{\frac{1}{2}}.$$

Substituting the value S_k into Eq. (8-19), we obtain

$$S = \frac{k}{k+1} \cdot 3{,}5 \cdot 10^{-6} \frac{F_{c1\,\text{эф}}}{r_{c1}^2 \beta^2} U_A^{\frac{1}{2}}. \tag{8-20}$$

Letting G stand for $2.33 \cdot 10^{-6} (F_{c1\,eff} / r_{c1}^2 \beta^2)$, we may write an expression for the mutual conductance in the form:

$$S = \frac{3}{2} \cdot \frac{k}{k+1} G U_{\text{A}}^{\frac{1}{2}}. \quad (8\text{-}20a)$$

Consequently, in tetrodes and pentodes, the mutual conductance is heavily dependent upon the current distribution, since the higher the current-distribution coefficient, the greater the transconductance will be, all other conditions remaining the same. Thus, in designing tetrodes and pentodes, we should attempt to decrease the screen-grid current.

The second static characteristic of tetrodes and pentodes is the interral resistance, defined as the ratio of a change in plate voltage to the change in plate current with the potentials on the remaining electrodes held constant:

$$R_i = \frac{\Delta U_a}{\Delta I_a}. \quad (8\text{-}21)$$

When plate voltage changes, there is a change in plate current, determined by two factors. First, a change in plate voltage leads to a change in the effective potential of the tube and, consequently, in the cathode current. The increment in cathode current is distributed between the plate and screen. Second, there is a redistribution of current between the plate and screen. The total change in plate current may be represented as a sum of two components:

$$\Delta I_a = \Delta I_{aD} + \Delta I_{a.k}. \quad (8\text{-}22)$$

The quantity ΔI_{aD} is the increment in plate current caused by a change in cathode current owing to a change in the effective potential of the tube. This quantity depends upon the penetrance of the tube, which determines the degree to which the plate voltage affects the current removed from the cathode. The increment $\Delta I_{a.k}$ is determined by the change in the distribution of currents with a change in plate voltage,

and depends upon the geometric relationships in the screen-plate space of the tube. Substituting into (8-21) the sum of these two increments in place of ΔI_a, we obtain:

$$R_i = \frac{\Delta U_a}{\Delta I_{aD} + \Delta I_{a.k}} \tag{8-23}$$

or

$$\frac{1}{R_i} = \frac{\Delta I_{aD}}{\Delta U_a} + \frac{\Delta I_{a.k}}{\Delta U_a}. \tag{8-23a}$$

Let $\Delta U_a/\Delta I_{aD} = R_{iD}$, and $\Delta U_a/\Delta I_{a.k} = R_{ik}$; then

$$\frac{1}{R_i} = \frac{1}{R_{iD}} + \frac{1}{R_{ik}}. \tag{8-24}$$

Consequently, the internal resistance of tetrodes and pentodes may be represented in the form of two resistances connected in parallel; R_{iD}, depending upon the penetrance of the tube, and R_{ik}, determined by the current distribution in the tube. If the penetrance of the tube is very low, a change in plate voltage has almost no effect upon the size of cathode current, and the resistance R_{iD} approaches infinity. In this case, it may be assumed that $R_i = R_{ik}$. If the penetrance is appreciable, and the influence of the plate voltage upon the cathode current cannot be neglected, then

$$R_i = \frac{R_{iD} R_{ik}}{R_{iD} + R_{ik}}. \tag{8-25}$$

The internal resistance may be found from the plate curve for the tube, with computation of the ratio $\Delta U_a/\Delta I_a$, as is shown in Fig. 8-14. It is clear that the flatter the plate curve, the less an increment in plate current will correspond to a given increment in plate voltage, and the higher will be the internal resistance of the tube.

Since pentodes generally have a higher internal resistance than do tetrodes, the plate curves for pentodes, with direct-interception operation, are flatter than the curves for tetrodes.

Plate curves for tetrodes and pentodes, taken with various voltages on the control grid, will not be parallel. The higher the control-grid voltage, the steeper the plate curve. This is explained by the fact that with a greater cathode current, there will be a greater change in plate current for precisely the same change in plate voltage. Consequently, the internal resistance drops with an increase in the control-grid voltage.

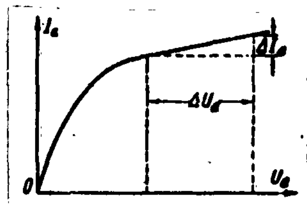

Fig. 8-14. Determining the internal resistance from the plate curve.

The amplification factor for tetrodes and pentodes is found from the ratio of the change in plate voltage to the change in control-grid voltage with the plate current and the voltages on the remaining electrodes constant:

$$\mu = \frac{\Delta U_a}{\Delta U_{c1}} \text{ where } I_a = \text{const.} \qquad (8\text{-}26)$$

In contrast to the triode, where a change in plate current can occur with the grid negative only as the result of a change in the cathode current, in the pentode, the changes in plate current owing to changes in plate voltage and to changes in control-grid voltage are different in character. For example, in order to maintain the plate current constant with an increase in the control-grid voltage, the plate voltage should be dropped. A rise in control-grid voltage, however, leads to an increase in the tube cathode current, while a drop in plate voltage causes a decrease in plate current owing to a redis-

tribution of currents between the plate and the screen grid with no substantial change in cathode current. Due to low penetrance, a drop in plate voltage sufficient to compensate a plate-current variation will not compensate a rise in the effective potential caused by an increase in control-grid voltage. Consequently, constant plate current is no guarantee of constant effective potential, as is the case with a triode.

Let us assume that in order to maintain the effective potential unchanged with an increase in control-grid voltage of ΔU_{c1}, it is necessary to decrease the plate voltage by an amount δU_a. From the condition that the effective potential must not change, it follows that

$$\Delta U_{c1} = D\delta U_a, \tag{8-27}$$

whence

$$\frac{\delta U_a}{\Delta U_{c1}} = \frac{1}{D}. \tag{8-28}$$

The quantity $1/D = \mu_0$ is called the electrostatic amplification factor of a tube. But in order to compensate a plate-current change with an increase in grid voltage, the plate voltage must change by an amount ΔU_a, less than δU_a, since a drop in plate current takes place basically owing to a redistribution of currents, and to a very slight degree owing to a drop in the effective potential. Consequently, the ratio

$$\frac{\Delta U_a}{\Delta U_{c1}} < \frac{\delta U_a}{\Delta U_{c1}} \ ; \ \mu < \mu_0,$$

i.e., the true ampliciation factor of a pentode (or tetrode) is less than the electrostatic amplification factor found on the basis of the penetrance.

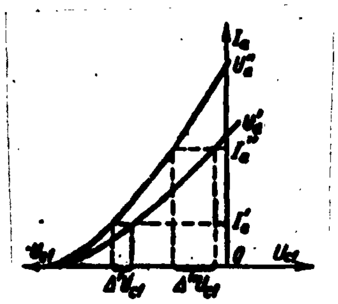

Fig. 8-15. Determining the gain of a pentode.

The true amplification factor of a pentode may be found from the family of plate-grid curves. Let us assume that there are two plate-grid curves, taken at plate voltages of U_a', U_a'' (Fig. 8-15). A change in plate voltage

$$\Delta U_a = U_a'' - U_a'$$

with the plate current I_a' constant corresponds to a change in control-grid voltage of $\Delta' U_{c1}$. The same change in plate voltage with a constant plate current I_a'' corresponds to an increment in grid voltage of $\Delta'' U_{c1}$. Since the plate-grid curves diverge, $\Delta'' U_{c1} > \Delta' U_{c1}$. Clearly, the amplification factor, defined as the ratio of the increment in plate voltage to the increment in control-grid voltage, will be different for different values of plate current. With a plate-current value of I_a, the amplification factor will be higher than for a greater plate current I_a''. As the potential on the control grid drops, the amplification factor rises, drawing near to the value of the electrostatic amplification factor at a plate current close to zero.

Thus, the amplification factor of pentodes and tetrodes depends upon the operating conditions of the tube, unlike the amplification factor for triodes, which is nearly independent of any change in

operating conditions.

As with triodes, the characteristics in tetrodes and pentodes may be measured by the three-readings method. In measuring tube characteristics, the change in grid voltage should be made as small as possible (on the order of 0.1-0.2 v). It is clear that the universal relation $SR_i = \mu$ holds for pentodes and tetrodes.

8-8. DYNAMIC CURVES AND CHARACTERISTICS FOR TETRODES AND PENTODES.

Owing to the advantages of beam-power tetrodes and pentodes in comparison with triodes that we have mentioned above, the former have received very wide application in voltage- and power-amplification circuits, and in many other radio and electronic circuits.

When these tubes are operated under dynamic conditions with a resistive load in the plate circuit, the same relationship is observed between the tube plate current and plate voltage as obtains in a triode operating under dynamic conditions:

$$i_a = \frac{E_a}{R_a} - \frac{1}{R_a} U_a,$$

i.e., the dynamic plate curve takes the form of a straight line.

Figure 8-16 shows a family of static plate curves for the 6P9 [6AG7] pentode; three dynamic curves are plotted, corresponding to a plate supply voltage of E_a = 300 v, and load resistances of 2, 5, and 15 kohm. As may be seen from the graph, the first dynamic curve intersects the static curves on the flat portions, corresponding to direct-interception operation, the second in part passes through bends in plate curves, while the third intersects plate curves taken for low grid voltages on flat portions, and plate curves taken for high U_{c1} on portions corresponding to electron-return conditions.

On the basis of the points of intersection of the dynamic curves

and the static curves, it is possible to plot the dynamic plate-grid

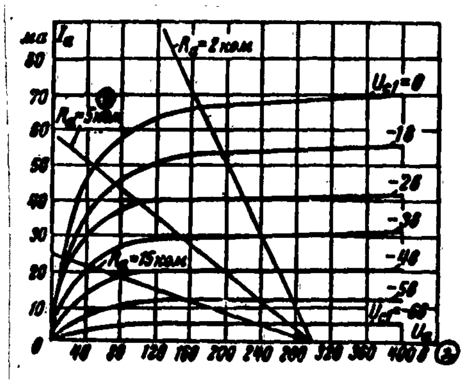

Fig. 8-16. Load lines in a family of plate curves for a 6P9 pentode.
1) kohm; 2) volts.

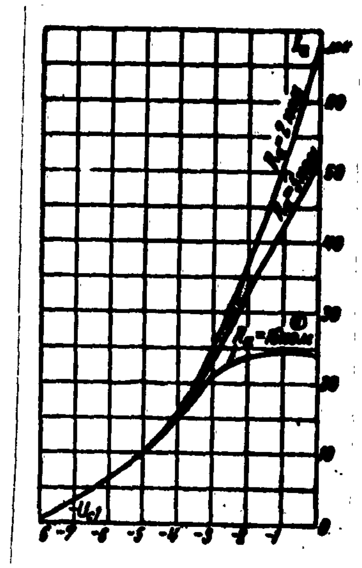

Fig. 8-17. Dynamic plate-grid (operating) curves for a 6P9 pentode.
1) kohm.

curves shown in Fig. 8-17. The dynamic plate-grid curve corresponding to a 2-kohm load resistance has its maximum slope in the control-grid interval from -3 v to 0 (below -4 v, all three curves coincide). The curve corresponding to a 5-kohm resistance has a somewhat less steep slope in the same voltage interval, but is more nearly linear. Finally, the curve plotted for a 15-kohm load resistance has a sharply decreasing slope in the -4 to 0 v control-grid range.

Consideration of these curves leads to the following conclusions.

First, size of the plate load should be so chosen that at the given amplitude of the alternating voltage on the control grid, the fluctuations in plate voltage will not go outside the plate-curve section corresponding to direct-interception operation. Operation of the tube under electron-return conditions causes considerable distortion of the amplified voltage, decreases the dynamic amplification factor, which is proportional to the transconductance, and, in addition,

leads to the dissipation of a large amount of power on the screen, which decreases circuit efficiency, and may cause the tube to fail, owing to overheating of the screen.

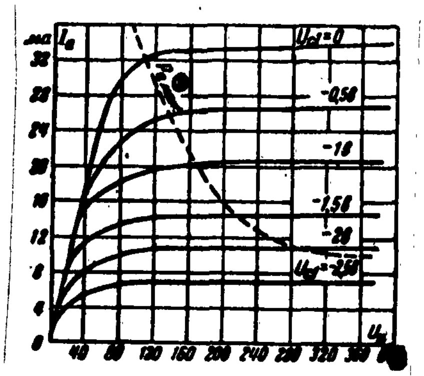

Fig. 8-18. Graph of the maximum permissible power for a family of static plate curves for a 6Zh4 pentode.
1) $P_{a\ max}$; 2) volts.

Second, with low load resistances, the plate-grid dynamic curve is nonlinear, which may cause nonlinear distortion of the voltage being amplified. The nonlinearity of a dynamic plate-grid curve is explained by the fact that the separations of static plate curves that are plotted for grid voltages changed by equal amounts are not equal; i.e., equal changes in control-grid voltage correspond to unequal changes in plate current. The most linear dynamic plate-grid curves are those corresponding to a value of plate load at which the dynamic plate-grid characteristic passes through sections where there is a bend in the static curves. Thus, where amplification with minimum distortion is required, the choice of plate load is confined to within rather narrow limits.

Since the family of plate-grid curves for tetrodes (and pentodes) that is confined to plate voltages corresponding to direct-interception operation forms a very narrow group of curves, the dynamic transconductance differs little in value from the static transconductance for the tube. The higher the internal resistance of the tube the

narrower the pattern of plate-grid curves, and, consequently, the less the dynamic and static transconductance differ. Actually, the dynamic transconductance equals, according to formula (7-14):

$$S_d = S\frac{R_i}{R_i + R_a}.$$

The magnitude of the load resistance R_a is normally considerably less than the internal resistance of the tube. The higher the value of R_i, the closer the ratio $R_i/(R_i + R_a)$ approaches one, and the closer together the values of the dynamic and static transconductances.

The dynamic amplification factor may be found with sufficient accuracy for practical calculations in terms of the product of the tube transconductance and the load resistance:

$$\mu_d \approx SR_a.$$

In choosing the value of the plate load resistance, it is necessary to be sure that the power dissipated on the plate of the tube under dynamic conditions does not exceed the permissible dissipation. To do this, a maximum-permissible-power curve is plotted on the family of plate curves for pentodes and tetrodes designed to operate in power-amplifier circuits (Fig. 8-18).

8-9. VARIABLE-MU TUBES.

Variable-mu tubes form a special class of high-frequency pentodes for voltage-amplification purposes. These tubes have appeared owing to the requirement that radio-receiving equipment operate satisfactorily both when receiving low-power signals from distant or weak transmitters, and when receiving powerful signals from nearby high-power transmitters. It is clear that the amplification given to signals received in RF stages of receiving equipment should be different

for different signal strengths. If this were not the case, an amplifier designed for strong-signal reception would not amplify weak signals enough for subsequent detection, and a high-gain weak-signal amplifier would amplify strong signals to such voltage levels that the signal received is distorted.

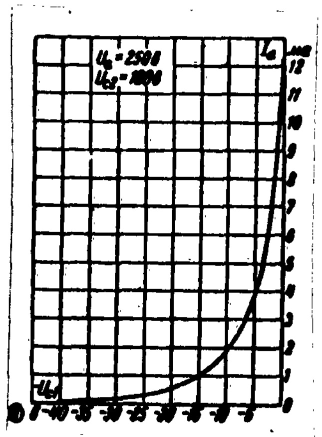

Fig. 8-19. Plate-grid curve for the 6K1P variable-mu RF pentode.
1) volts.

The problem of obtaining RF voltage gain as a function of the signal voltage at the input of the amplifier is solved by utilizing pentodes whose plate-grid curves have the shape shown in Fig. 8-19. In order to obtain such a curve, the control grid of the pentode is made with varying spacing, in which the majority of the turns are close together, and one or two turns are spaced far apart (they are usually in the central portion of the grid) as shown in Fig. 8-20. Such a tube may take the form of two tubes connected in parallel, one having a long, closely spaced grid, and the other, a short, widely spaced grid. Since the transconductance is proportional to the working length of grid, the curve for the first section, which is greater in length, will have a correspondingly greater transconductance. The grid penetrance will be low for this section, and cutoff will occur at a low absolute

value of the negative grid potential. Consequently, the first section will match the right-hand curve, with a high transconductance.

The second section of the tube, with a short grid, will have high penetrance, and, consequently, the left-hand curve, with low transconductance. The curves for both sections of the tube are plotted on the graph shown in Fig. 8-21 (dashed lines). The total tube current equals the sum of the currents of both sections, and thus, combining the curves, we obtain the resultant plate-grid curve consisting of two parts: a gently rising portion extending for some distance along the control-grid voltage axis, and a part that is very steep and corresponds to a small range of grid-voltage changes. Depending upon the bias on the control grid, the operating point may be on either the

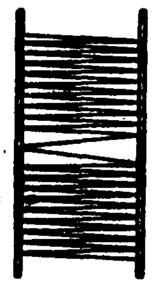

Fig. 8-20. Grid of a variable-mu tube.

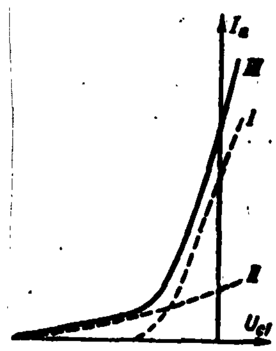

Fig. 8-21. Plate-grid curve for a variable-mu tube. I) Curve for section with small grid spacing; II) curve for section with wide grid spacing; III) resultant tube curve.

flat or the steep portion of the characteristic; in the first case, the tube will have low gain for a strong signal, in the second case, high gain for a weak signal. The bias is varied automatically in receiving equipment, depending upon the amplitude of the received-signal voltage, using a special automatic sensitivity-control circuit (ARCh). In this circuit, the amplified signal is rectified, and the

rectified voltage used as the bias voltage.

Variable-mu tubes are widely employed in radio receiving equipment. It should be noted that if the control grid of such a tube is made with a continuously varying pitch, it is possible to obtain a plate-current — grid-voltage dependence that does not follow the three-halves law, but a desired law; this is made use of for certain special purposes.

Chapter Nine

TUBES FOR FREQUENCY CONVERSION

9-1. TUBES WITH DUAL PLATE-CURRENT CONTROL.

In radio-receiving devices, and in particular, in equipment designed for receiving HF and VHF signals, so-called superheterodyne circuits are widely used. The operation of such circuits is based upon the "mixing" of two voltages that differ from each other in frequency in such manner that there is a constant difference between the frequencies being mixed.

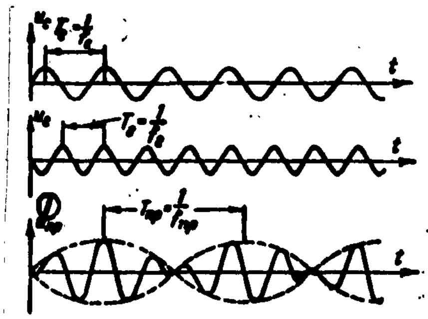

Fig. 9-1. Formation of beats as the result of combining two oscillations with different frequencies.
1) u_{pr}.

The principle of frequency mixing is illustrated graphically in Fig. 9-1. Upon combining a voltage u_1 with a frequency f_1, and a voltage u_2 with a frequency f_2, we get a resultant voltage that has a periodically varying amplitude. Oscillations of this type are called beats. Mathematical analysis shows that the amplitude-variation frequency (beat frequency) equals the difference in the frequencies of the voltages combined:

$$f_б = f_1 - f_1.^{*} \qquad (9\text{-}1)$$

In a radio-receiving device, the voltages u_1 and u_2 are the signal voltage, and a variable voltage at an intermediate frequency, created in the receiving device itself by a special oscillator — the local oscillator. This oscillator is tuned together with the tuned input (antenna) circuit of the receiver so that the difference between the local-oscillator frequency and the frequency of the signal remains constant, regardless of the tuning of the receiver.

For frequency "mixing," tubes are used in which it is possible to use two voltages to control the plate current. Such tubes are called mixers. In contrast to the tubes previously discussed, having a single control electrode — the first grid — mixers have two control electrodes.

A pentode may be used as a very simple mixer, with the first and third grids utilized as control elements. With standard pentodes, however, the controlling action of the third is slight, owing to the large grid spacing. Thus, mixer pentodes (dual-control pentodes) are constructed with control grids having low penetrance.

A negative bias voltage is applied to the third grid of a pentode mixer. As a result, an electric field that strongly retards electrons is set up in the space between the screen and suppressor grids, and a retarding space charge forms, acting as a virtual cathode. This virtual cathode, together with the suppressor and plate of the tube form a triode-type system, whose plate current can be varied by varying the potential differences between the suppressor grid and the

* [б - b - biyeniye - beat.]

potential minimum set up by the space charge.

The density of the space charge, and, consequently, of the potential in the space between the screen and suppressor grids as well, changes at a frequency equal to the frequency of the signal applied to the control grid. The magnitude of the current drawn from the virtual cathode to the plate changes, first, under the action of the varying signal voltage on the suppressor grid and, second, as a result of the change in potential of the virtual cathode taking place at the frequency of the signal applied to the control grid. Thus, the plate current changes owing to the simultaneous influence of both of the frequencies being mixed, and oscillations of the beat type appear in the plate circuit.

Upon transition to the positive-potential region of the supressor grid, the virtual cathode is dissipated, return of electrons to the screen grid ceases, and a further increase in the potential of the suppressor grid has little effect upon the magnitude of the plate current. Thus, the curve showing the dependence of the plate current upon the suppressor-grid voltage with constant voltages on the remaining electrodes is quite steep in the negative-voltage regions, and becomes flat when the suppressor goes positive.

Figure 9-2 shows curves of $I_a = f(U_{c3})$ for a type 6Zh2B (REGE) dual-control pentode, taken with various control-grid voltages. A change in suppressor-grid potential in the negative region causes a considerable change in plate current and, consequently, the suppressor may be utilized as a second control electrode for the tube.

Since a change in suppressor-grid potential causes a change in the current distribution, the mutual conductance for the control grid is also a function of the suppressor-grid potential. The quantitative relationship between the mutual conductance and the suppressor-grid

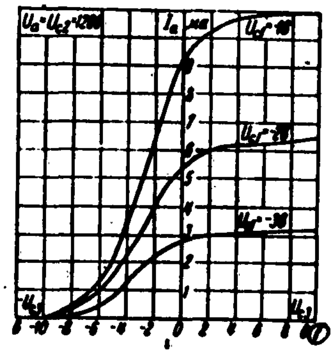

Fig. 9-2. Grid-plate characteristics for suppressor of 6Zh2B pentode.
1) Volts.

potential of a pentode with dual control may be found by plotting a family of grid-plate curves, taken at several values of suppressor-grid voltage. On the basis of this family, it is not difficult to plot a graph of $S_1 = f(U_{c3})$ (Fig. 9-3). This function is nearly linear in the region of negative suppressor-grid voltages.

In order to determine the relationship between the mutual conductance and the suppressor-grid voltage, we introduce the so-called conversion factor, which indicates the change in mutual conductance for each volt of change in suppressor-grid voltage:

$$K_{np} = \frac{\Delta S_1}{\Delta U_{c3}} \; [\text{ма}/в^2].^{*} \qquad (9\text{-}2)$$

It is clear that the greater the controlling action of the suppressor grid, the greater the conversion factor will turn out to be.

Figure 9-4 shows a circuit for the utilization of a pentode mixer. The signal voltage is applied to the tube control grid from an input tuned circuit

---

* [в - v - vol't - volt.]

$$u_c = U_{mc} \sin \omega_c t,$$

while the local-oscillator voltage goes to the third grid:

$$u_r = U_{mr} \sin \omega_r t. \quad *$$

Since the mutual conductance of a pentode changes as the suppressor voltage changes, where there is a varying voltage on the suppressor, the mutual conductance will also become a varying quantity. If at the operating point the mutual conductance has a value S_0 (static mutual conductance with nominal voltages on all electrodes), then with a varying voltage on the suppressor grid, the mutual conductance will depart from S_0 by an amount $K_{pr}u_c$. The instantaneous value of the mutual conductance may be expressed in terms of the following equation

$$s = S_0 + K_{np}u_r = S_0 + K_{np}U_{mr} \sin \omega_r t. \qquad (9\text{-}3)$$

Since an alternating signal voltage is applied to the control grid of the pentode, an alternating current will flow in the plate circuit equal to:

$$i_a = su_c = (S_0 + K_{np}U_{mr} \sin \omega_r t)U_{mc} \sin \omega_c t. \qquad (9\text{-}4)$$

Removing parentheses, we obtain:

$$i_a = S_0U_{mc} \sin \omega_c t + K_{np}U_{mc}U_{mr} \sin \omega_c t \sin \omega_r t. \qquad (9\text{-}4a)$$

Carrying through trigonometric manipulations in Eq. (9-4a), we obtain an expression for the alternating component of the plate current:

---

* [u_r - u_g - $u_{geterodin}$ - $u_{local\ oscillator}$.]

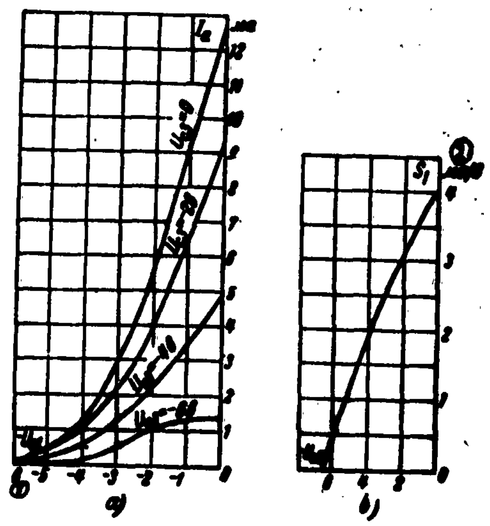

Fig. 9-3. Mutual conductance with respect to first grid as a function of suppressor-grid voltage for a 6Zh2B pentode.
1) Volts; 2) ma/v.

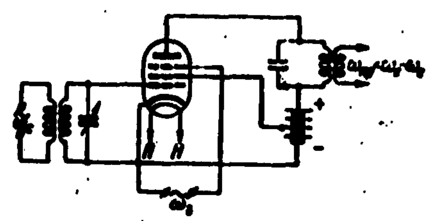

Fig. 9-4. Circuit for the utilization of a pentode mixer.

$$i_a = S_0 U_{mc} \sin \omega_c t + \frac{1}{2} K_{\text{пр}} U_{mc} U_{mr} \times$$
$$\times \cos(\omega_r - \omega_c) t - \frac{1}{2} K_{\text{пр}} U_{mc} U_{mr} \cos(\omega_r + \omega_c) t. \qquad (9\text{-}5)$$

Consequently, the alternating plate current takes the form of a sum of three sinusoidal currents, one of which has a frequency equal to the difference of the local-oscillator and signal frequencies.

A tuned circuit is connected into the plate circuit of the tube, and is tuned to the intermediate frequency, equal to the difference in the frequencies of the local-oscillator and signal. As a result, a component is separated out of the alternating plate current that has a frequency equal to the intermediate frequency. The power developed in

the tuned circuit by the remaining plate-current components is slight.

Since the local-oscillator tank circuit is tuned together with the input tuned circuit, the difference in frequency between the signal and the local oscillator remains fixed, and the tuning of all stages in the receiver following the mixer remains fixed, whatever the tuning of the input circuit.

For frequency conversion, it is desirable to have the maximum amplitude of the alternating plate-current component at the intermediate frequency. Equation (9-5) shows that the amplitude of this component will equal:

$$I_{m\,\text{пр}} = \frac{1}{2} K_{\text{пр}} U_{mc} U_{mr}. \tag{9-6}$$

The quality of a tube in the role of a mixer is determined by the conversion transconductance which shows what amplitude of the intermediate-frequency alternating component may be obtained with a signal-voltage amplitude of 1 v for a given amplitude of the local-oscillator voltage:

$$S_{\text{пр}} = \frac{I_{m\,\text{пр}}}{U_{mc}} = \frac{1}{2} K_{\text{пр}} U_{mr}. \tag{9-7}$$

The conversion transconductance is proportional to the conversion factor and, consequently, depends upon the tube characteristics.

Dual-control pentodes are widely employed in so-called coincidence circuits of various types of computing devices and automatic devices. These circuits utilize the property of these pentodes that makes it possible to cut them off by means of a voltage upon either the first or third grid. As a result, the appearance of plate current from the tube is possible only when positive signals are applied simultaneously to both controling grids.

As mixer tubes, these pentodes possess several drawbacks; the

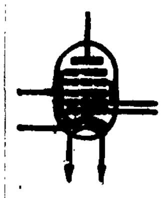

Fig. 9-6. Heptode.

9-2. MULTIGRID FREQUENCY-CONVERSION TUBES.

The utilization of mixer tubes in a frequency-conversion stage of a receiver requires the presence of an electron tube used in an auxiliary-frequency oscillator circuit: this is the local oscillator. Circuits with separate mixer and local-oscillator tubes are seldom used in practical receiving devices. In most cases, the functions of mixer and local-oscillator tube are combined in a single tube, called a frequency converter. Triode-hexodes, triode-heptodes, and heptode converters (pentagrid converters) may be used as frequency-conversion tubes.

Let us consider a frequency-conversion circuit using a pentagrid converter (Fig. 9-7). The cathode, first grid, and second grid form a triode, connected into a Hartley oscillator circuit. The stream of electrons, modulated by the local-oscillator signal frequency passes into the space between the second grid and the third grid which acts as the second control grid. The signal voltage is applied to the third grid, and the intermediate frequency appears in the plate circuit, as in the case of the pentode mixer discussed above.

The drawback to the pentagrid converter is the possibility of the appearance of electronic coupling between the local-oscillator and receiver-input circuits due to the electrons returning from the virtual cathode to the local-oscillator grid. This phenomenon may lead to a change in local-oscillator frequency. A decrease in electronic coupling

basic fault is the large capacitance between the suppressor and plate, which limits the utilization of such tubes as mixers when the operating frequencies become high. Better mixer tubes are those having four and five grids — hexodes and heptodes.

The hexode (Fig. 9-5) has four grids, of which the first and third are control grids, and the second and fourth are screen grids operating at high positive potentials. The screen voltage in hexode mixers generally amounts to 40 to 60 per cent of U_a.

A dynatron effect can appear when a hexode is operated in a mixer circuit associated with a high-amplitude alternating plate-voltage component. Heptode mixers are free from this fault; these tubes have a fifth grid in the space between the fourth grid and the plate — a suppressor grid connected to the cathode (Fig. 9-6).

The basic characteristics of mixer tubes are the first-grid mutual conductance S_1, the third-grid transconductance S_3, and the internal resistance R_i. The first two parameters determine the possibility of obtaining the required conversion transconductance from the mixer. Especially important is high first-grid mutual conductance, which ensures conversion of relatively weak signals applied to the first grid of the tube. An increase in tube internal resistance makes it possible to utilize tuned IF circuits, and to obtain greater dynamic gain from a mixer stage. It is clear that hexodes have greater internal resistance than do pentode mixers; thus, use of hexodes in mixer stages is more advantageous. Heptode mixers have still higher internal resistance.

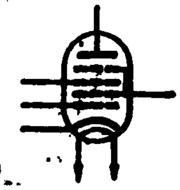

Fig. 9-5. Hexode.

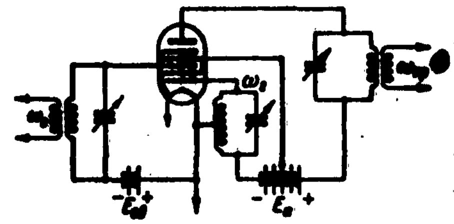

Fig. 9-7. Frequency-conversion circuit using pentagrid converter.
1) ω_{pr}.

between the signal and local-oscillator circuits is achieved in special pentagrid-converter designs that provide for the interception of the returning electrons by the second grid.

In order to avoid the return of electrons to the region of the local-oscillator grid in a pentagrid converter, it is possible to utilize the design shown schematically in Fig. 9-8. Owing to the presence of four traverses of considerable diameter at the third grid, the stream of electrons is focused into four high-density beams. The second (screen) grid is made solid around the traverses. The electrons reflected from the third grid strike the solid portion of the second grid, and do not penetrate into the region of the tube that serves as the local-oscillator triode.

Also finding application in frequency-conversion stages of receiving devices are so-called triode-hexodes, which take the form of two tubes combined in a single envelope: a triode and a hexode. The triode section of the tube operates in a local-oscillator circuit, with the triode grid connected within the tube to one of the control grids of the hexode – the first or the third (Fig. 9-9).

Still better is the triode-heptode design (Fig. 9-10). In this tube, the heptode, with a fifth, screen grid connected with the cathode and the metal shield, is used as the mixer. The second and fourth

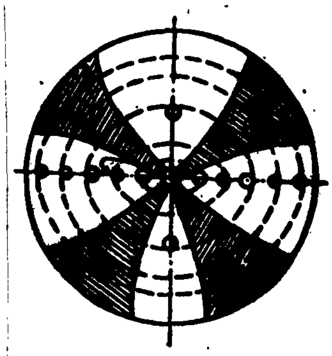

Fig. 9-8. Schematic representation of the structure of a pentagrid converter.

grids of the heptode section are screen grids, and are interconnected within the tube. The triode grid has a separate lead which makes it possible to use the tube in various types of circuits.

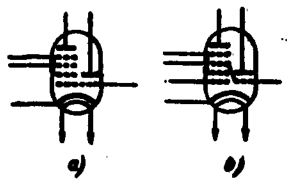

Fig. 9-9. Symbols for triode-hexodes. a) With a common control grid; b) with jumper between triode grid and third grid of hexode.

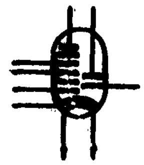

Fig. 9-10. Symbol for triode-heptode.

In mixer and converter tubes, the grid to which the signal voltage is applied is frequently made with variable spacing so as to provide a variable-mu characteristic at this grid.

The basic characteristics of converter tubes are the first-grid mutual conductance, normally called the local-oscillator transconductance, the internal resistance of the tube, and the conversion transconductance, defined, as with mixer tubes, as:

$$S_{pr} = K_{pr}U_{mg}/2.$$

Chapter Ten

RECEIVING-AMPLIFIER TUBES

10-1. EXTERNAL APPEARANCE OF RECEIVER-AMPLIFYING TUBES.

Existing receiving-amplifying tube designs may be grouped on the basis of their external appearance as follows:

1. Tubes in glass envelopes (Fig. 10-1a, b).
2. Tubes in metal envelopes (Fig. 10-1c, d).
3. Tubes in glass envelopes with an external metal shield and a lock in the base locating pin (Fig. 10-1e).
4. Miniature tubes with cylindrical glass envelopes and flat mounts (Fig. 10-1f, g).
5. Subminiature tubes (Fig. 10-1h, i).

Receiving-amplifying tube designations are determined by their type and external appearance, and consist of four elements: the first element is a number, the second a letter, the third is a second number, and the fourth, a second letter. The first number in the designation of a tube indicates the approximate filament voltage (for example, the

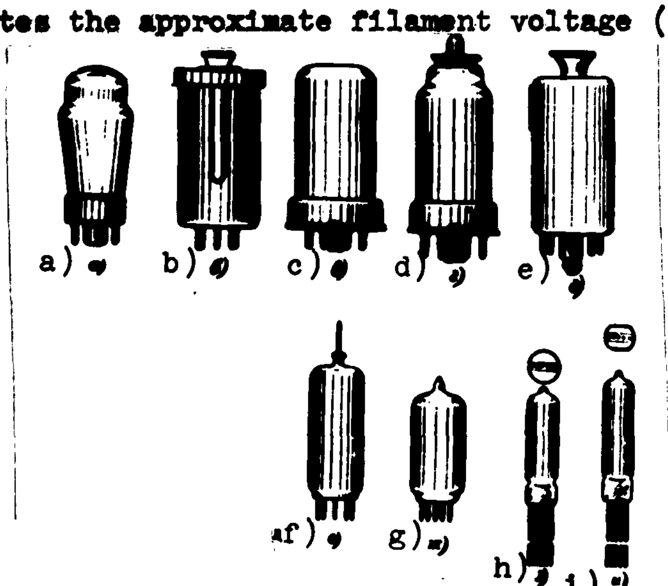

Fig. 10-1. External view of receiving-amplifying tubes.

number 6 indicates that the tube filament voltage equals 6.3 v, the

W-16

number 2 is used for a tube with a filament voltage of 2 v, 2.2 v, etc.). The letter indicates the type of tube: D — diode, Kh — dual diode, Ts — kenotron rectifier, S — triode, E — tetrode, P — audio-frequency power pentodes and beam tetrodes, Zh — RF pentodes and beam tetrodes with short plate-grid characteristic curves, K — RF pentodes and beam tetrodes of the variable-mu type (with an extended set of characteristic curves), and A — frequency-conversion tubes (hexodes and heptodes) with two control grids. Where several types of tubes are combined in a single envelope, the designations used are: G — triode with one or two diodes, B — pentode with one or two diodes, N — twin triode, F — triode with pentode, I — triode with multigrid mixer (hexode or heptode), and E — tuning indicator.

The second number in the designation indicates the sequence number of the given type of tube, and the final letter defines the external tube construction: S — glass, L — glass tube with external metal shell with lock in base locating pin, P — miniature (small-button), A — subminiature with an envelope diameter of 10-11 mm, Zh — "acorn" type tube, and D — lighthouse tube.* Designations for tubes in metal envelopes consist of three elements, with no final letter. Thus, the 6Zh1P is a RF pentode with a short characteristic curve, miniature, with a filament voltage of 6.3 v, and the 12K4 is an RF pentode, variable-mu, in a metal shell, with a 12.6-v filament voltage.

10-2. ADDITIONAL TUBE CHARACTERISTICS.

In addition to the tube characteristics that have been discussed, the possibility of application of a tube to various circuits is to a large degree determined by several other properties. These properties basically determine the nature and degree of influence of several

---

* For a detailed discussion of "acorn" and lighthouse tubes, see Chapter Twelve.

processes occurring in tubes.

a) Interelectrode insulation. In the description of the passage of current through electron tubes, conductance owing to the insulators separating tube electrodes has been assumed to be zero; it has also been assumed that the magnitude of the electrode currents is determined solely by the electrostatic fields depending upon the electrode potentials, and the space charge filling the space between the electrodes.

In practice, in addition to currents carried by electrical charges through a vacuum-filled gap between electrodes at different potentials, there also exist leakage currents appearing owing to insulation imperfections. A deterioration in the properties of insulators utilized in receiving-amplifying tubes (glass, mica, and ceramics) is caused either by the presence in the insulators of various conducting impurities, or by the formation of conducting films on the surface of the insulators while the tube is being processed and operated. Such films may be formed by particles of the getter, as well as by dust particles from the cathode and other heated tube elements.

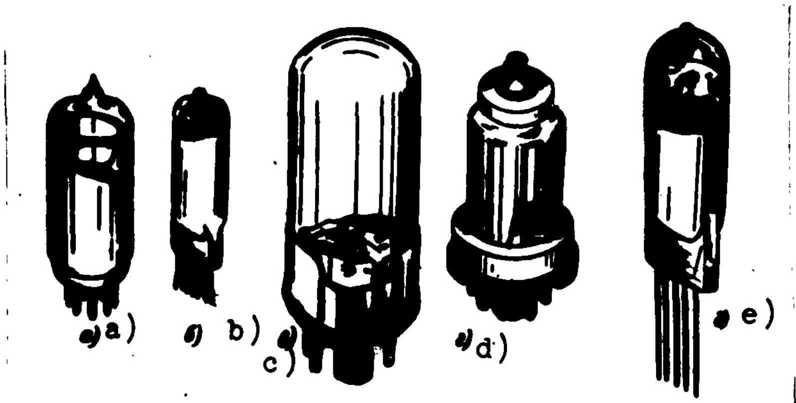

Fig. 10-2. Arrangement of getters in receiving-amplifying tubes.

The appearance of leakage currents may substantially impair the operation of a vacuum-tube circuit, and may in certain cases lead to failure of the tube. Thus, in receiving-amplifying tubes, there is

61-10~

decreasing grid temperature and, second, applying a surface coating to the grids that will prevent a drop in work function when barium from the cathode is deposited on the grids.

In order to decrease the operating temperature of control grids, their traverses are frequently manufactured of copper wire, which increases the amount of heat carried off. Carbonized radiators are welded to the end of the grid traverses, increasing the heat carried off owing to radiation into space.

Precious metals (platinum, gold, silver) are normally used as protective coatings. The best of these is gold which, first, is easily applied to a wire pulled through a molten drop of gold in a protective atmosphere and, second, goes into solution with free barium, which results, as it were, in "poisoning" of the newly forming oxide cathode. The utilization of platinum is complicated for technological reasons, while the utilization of silver should be avoided, since it evaporates easily, which leads to a deterioration in the interelectrode insulation.

The appearance of thermionic emission as a result of increased grid temperature sometimes is caused by the presence of leakage currents between the grid and other electrodes, which decreases the bias on the grid, leading to an increase in the power dissipated in the tube, and to additional heating up of the grid.

A third reason for the appearance of inverse grid current is the ion current that occurs when the vacuum in the tube is not high enough. The causes for the appearance of an ion current are insufficiently careful outgassing of parts during evacuation of the tube, and a poor-quality getter.

In the majority of receiving-amplifier tubes, the total inverse grid current should not exceed 0.5-3 μa under working conditions, and in certain tubes 0.1-0.2 μa, to insure normal operation of the tubes in

normally a predetermined magnitude of minimum insulation resistance (or maximum leakage current) between electrodes. For standard types of receiving-amplifying tubes, the insulation resistance is on the order of 10^8-10^{10} ohms.

In order to provide the required value of insulation resistance, several designs and process methods are employed. The getter is so located that particles from it are deposited on a portion of the envelope surface where there are no electrode leads, and so that they do not strike mica or ceramic insulators (Fig. 10-2).

In some cases, parts of a tube are protected against getter particles by a special mica shield. Materials chosen for use in manufacturing tube parts, and in particular parts operating at high temperatures, should have a low rate of evaporation. This requirement also applies to metals utilized for coatings on tube parts. Alkali and alkali-earth metals are very volatile. Following, in order of decreasing rate of evaporation, are chromium, silver, aluminum, copper, and gold.

In order to decrease surface conductance appearing as a result of deposition of particles on the surface of mica, forming conducting films, the surface of mica insulators is coated with a coarse layer of magnesium oxide (MgO) or alundum (Al_2O_3). Very coarse insulating coatings, first, increase the length of the electrical path between electrodes along the insulation and, second, make a poorer contact between the particles of conducting films. Additional apertures are sometimes cut out in order to provide still longer leakage-current paths along insulator surfaces between openings for mounting electrodes.

In tubes with indirectly heated cathodes, in which the heater is insulated from the cathode tube, the cathode-heater resistance is of great importance. In certain circuits, the voltage between cathode and heater may reach hundreds of volts, which would cause considerable

leakage currents where insulation was poor. The basic causes of deterioration in heater-cathode insulation are, first, contamination of the alundum covering the heater and, second, porousness and cracks in the alundum coating itself. Since the heater operates at quite a high temperature, a thermionic-emission current may appear, which in the presence of cracks in the alundum layer will flow through the cathode-heater circuit. Possible causes of increased conductance in an alundum layer are diffusion of tungsten from the base-layer metal of the heater to the alundum layer, and the formation of aluminum tungstate as a result of electrolysis of the alundum.

In certain types of tubes designed for operation in high-voltage circuits with a high voltage between cathode and heater, the heaters are not coated with alundum, but are located within channels inside a ceramic tube which is in turn inserted within the cathode tube. Such tubes may operate with cathode-heater voltage differences amounting to several thousand volts.

The magnitude of the leakage currents between cathode and heater in the majority of types of receiving-amplifying tubes does not exceed 20-50μ a with cathode-heater voltage differences on the order of 100-150 v.

b) Inverse grid currents. When small negative potentials are applied to a grid, a small electronic current is observed to flow from the cathode to the grid owing to the initial velocities of the electrons and the contact potentials between grid and cathode. At a specific grid potential, this current ceases, since the energy of the thermal electrons proves inadequate for overcoming the retarding field of the grid. This potential is called the grid-current cutoff potential. Its value depends upon the temperature and emission properties of the cathode, and also upon the grid-cathode contact potential.

In addition to the electronic current from cathode to grid at negative grid potentials, in the majority of tubes there is a current in the reverse direction called the inverse grid current. The reasons for the appearance of inverse grid current are thermionic emission from the grid, ionization of residual gases in the tube, and imperfections in the insulation.

Control grid temperatures, even in receiving-amplifying tubes delivering heavy power, do not exceed 400-450°C. At such temperatures, the thermionic emission from the surfaces of pure metals is so small that it has no noticeable effect upon tube operation. During the operation of a tube with an oxide-coated cathode, however, evaporation of barium and its oxides from the cathode takes place. As a result, a layer forms on the grid that resembles in structure the surface of an oxide-coated cathode (a mixture of pure metallic barium with barium oxide), which leads to a considerable drop in the work function of the grid surface, and to the appearance of thermionic emission.

The appearance of thermionic emission from the grid is especially dangerous when a tube is operated in circuits that have a high resistance in the grid circuit. Thermionic-emission current flows in the cathode-grid and plate-grid circuit. When current appears in the grid circuit, a portion of the voltage applied to the grid appears across this resistance as a voltage drop, resulting in the grid becoming more positive. This leads to an increase in plate current and, consequently, in the power dissipated on the plate, and to a rise in plate temperature, which in turn causes additional overheating of the grid. An increase in grid temperature produces a rise in grid thermionic emission. Thus, a snowballing increase in plate current develops, leading to overheating of the tube and to tube failure.

Ways of decreasing grid thermionic emission are, in the first place,

circuits.

c) Plate-current cutoff voltage. In many vacuum-tube circuits, for example, in so-called relaxation oscillators and generators of special wave forms, the plate-current cutoff voltage (blocking voltage) assumes great importance. As has been said previously, the theoretical value of cutoff voltage depends solely upon the penetrance of the grids. In practice, the absolute value of cutoff voltage turns out to be larger. The reasons for this are, first, nonuniformities in grid spacing and grid-cathode separation, which amounts to nonuniformity in the value of the amplification factor and, second, the presence of cathode-plate leakage. In addition, the initial velocities of electrons have a considerable effect upon the value of cathode current on the initial section of the plate-grid curve. As a result, the initial section of the curve for real tubes differs noticeably from that present in theoretical curves.

The characteristic determining the degree of deviation of the initial section of the curve from the theoretical curve (lengthening of the "tail" of the plate-grid curve), in receiving-amplifying tubes may be either the maximum permissible value of plate current at a given negative grid potential (current on the "tail" of the characteristic), or the grid voltage at which the plate current does not exceed a given value.

d) Internal tube noise. There is always an alternating voltage component across the load resistor of a vacuum-tube amplifier, even in the absence of input signal and even where theoretically pure DC voltages are delivered by the power supply. The causes for the appearance of this alternating component are random (fluctuating) variations in the currents flowing through the resistors and tubes of the amplifier, called noise. When a weak signal is applied to the input of an amplifier,

the alternating component across the load, obtained as a result of amplifying the signal, may prove to be on the same order as the random voltage fluctuations. In this case, the amplified signal may be distorted or even "drowned out."

The nature of the fluctuating current changes in resistances and tubes is associated with the possibility of random changes in the velocities and directions of motion of the current carriers — the electrons. In resistances, the random motion of electrons owing to their own velocities is superimposed upon the direction of electron motion effected by the electric field. As a result, at each instant of time, the concentration of electrons in individual volume elements of the resistor will be different, which leads to the appearance of equalizing currents between these elements. The superposition of the equalizing currents upon the current created by the applied potential difference leads to the appearance of fluctuations in the total current through the resistance. The current fluctuations create an alternating voltage drop — a noise voltage — across resistances.

Investigation of resistance noise has shown that noise voltage may be resolved into a series of sinusoidal harmonic components, covering the frequency spectrum from the low audio to the super-high, and the amplitude of all harmonics are about the same. The theory gives the following expression for the mean-square noise voltage in a given frequency band Δf:

$$E_{n}^{2} = 4kTR\Delta f,^{*} \qquad (10\text{-}1)$$

where T is the temperature of the resistance, $^{\circ}$K; R is the size of the resistance, ohms; k is the Boltzmann constant.

---

* [E_{n} - E_{sh} - E_{shum} - E_{noise}.]

In electron tubes, the appearance of noise is primarily associated with so-called shot noise, which is essentially explained by the fact that the number of electrons emitted by a cathode, and their initial velocities, are different at different instants of time. When a tube is operated at saturation, the plate-current fluctuations duplicate the cathode-emission current fluctuations. The majority of tubes, however, operate under space-charge-limited-current conditions. The presence of the space charge near the cathode attenuates the effects of emission-current fluctuations, and the magnitude of the noise drops considerably.

Of the electron tubes with grids, triodes have the least noise; the appearance of noise in triodes is chiefly determined by the weakened shot noise. Additional noise sources in multigrid tubes are the instability of secondary emission from tube electrodes, and fluctuations in current distribution associated with the fact that the directions of the velocities of individual groups of electrons are not constant with time.

Where the vacuum in a tube is not high enough, an additional noise component is created owing to ionization of residual gas. The electrons appearing during ionization flow to the plate, increasing the plate current. Positive ions, moving toward the cathode fall into the space-charge region, and partially neutralize the space charge. The instant the ions arrive in the space-charge region, the potential of this section of the space rises sharply, and there is a corresponding increase in the current from the cathode to this area. The magnitude of the noise appearing upon deterioration of the vacuum in a tube is the larger, the larger the ion current.

All of the factors enumerated above create noise that may be resolved into a series of harmonic components covering approximately the same frequency spectrum as does resistance noise. Additional

In electron tubes, the appearance of noise is primarily associated with so-called shot noise, which is essentially explained by the fact that the number of electrons emitted by a cathode, and their initial velocities, are different at different instants of time. When a tube is operated at saturation, the plate-current fluctuations duplicate the cathode-emission current fluctuations. The majority of tubes, however, operate under space-charge-limited-current conditions. The presence of the space charge near the cathode attenuates the effects of emission-current fluctuations, and the magnitude of the noise drops considerably.

Of the electron tubes with grids, triodes have the least noise; the appearance of noise in triodes is chiefly determined by the weakened shot noise. Additional noise sources in multigrid tubes are the instability of secondary emission from tube electrodes, and fluctuations in current distribution associated with the fact that the directions of the velocities of individual groups of electrons are not constant with time.

Where the vacuum in a tube is not high enough, an additional noise component is created owing to ionization of residual gas. The electrons appearing during ionization flow to the plate, increasing the plate current. Positive ions, moving toward the cathode fall into the space-charge region, and partially neutralize the space charge. The instant the ions arrive in the space-charge region, the potential of this section of the space rises sharply, and there is a corresponding increase in the current from the cathode to this area. The magnitude of the noise appearing upon deterioration of the vacuum in a tube is the larger, the larger the ion current.

All of the factors enumerated above create noise that may be resolved into a series of harmonic components covering approximately the same frequency spectrum as does resistance noise. Additional

noise is created as a result of the so-called flicker effect which consists of a continuous change in the activity of individual sections of the cathode, with a corresponding change in the emission from the sections. In contrast to all of the factors enumerated previously that are responsible for the appearance of noise, the flicker effect creates fluctuations only at low frequency. The level of the low frequency noise appearing as a result of the flicker effect depends basically upon the quality of cathode treatment during production.

In order to evaluate the noise level of a tube, it is possible to replace the "noisy" tube with an equivalent imaginary "noiseless" tube having a resistance connected into the grid circuit, the noise voltage across which creates the same plate-current fluctuations in the tube. Evidently, the greater the noise level in the tube, the greater the value of the equivalent grid noise resistance.

The equivalent noise resistance may be calculated from the following approximate formulas:

$$R_{э.ш}=\frac{2,5}{S}\ [\mathrm{kohm}]\ * \qquad (10\text{-}2)$$

for triodes and

$$R_{э.ш}=\frac{I_a}{I_a+I_{c2}}\left(\frac{2,5}{S}+\frac{20I_{c2}}{S^2}\right)\ [\mathrm{kohm}] \qquad (10\text{-}3)$$

for pentodes. In both formulas, S is the transconductance of the tube in ma/v.

The value found for the equivalent noise resistance corresponds to the value of effective noise voltage determined from formula (10-1), on the basis of which it is possible to determine the minimum permissible value of signal voltage. In order to keep tube noise from

---

* [$R_{э.ш}$ - $R_{e.\ sh}$ - $R_{ekvivalentnyy\ shum}$ - $R_{noise\ equivalent}$.]

drowning out the amplified signal, the signal voltage at the input of the tube should be considerably greater than the noise voltage. Consequently, the relative influence of noise is decreased as the tube transconductance rises, and at a high transconductance, it is possible to use low signal voltage, or a wide amplifier frequency range.

e) Mechanical strength and stability of tubes. During operation of a tube, it is possible to test for mechanical effects of various types: shocks, vibration, shaking, etc. These may adversely affect the strength of mountings for parts of the tubes, and the mutual location of tube parts. In addition, during the application of loads, elastic deformations and oscillation of parts may occur, causing a change in tube characteristics.

Of the types of mechanical effects upon tubes enumerated, vibration is one of the most frequently encountered. During oscillation of tube parts caused by vibrations, continuous changes in tube characteristics occur, owing to which an AC component will appear in the plate circuit, causing an alternating voltage drop across the plate load, called a vibration-noise voltage. Mechanical oscillations of the control grid and cathode have the greatest effect upon the magnitude of vibration noise; thus, in vibration-resistant tubes, special attention should be paid to the mounting of these elements.

The characteristic normally given as a measure of tube vibration resistance is the magnitude of the vibration-noise voltage arising with a given load resistance in the plate circuit under vibration conditions. The vibration endurance of tubes is characterized by the changes in their characteristics that take place as a result of a rather extended period of vibration; here, after being subjected to extended vibration, the tube is measured for its basic static characteristics and vibration-noise voltage.

The endurance of tubes when subjected to other types of mechanical action (shocks, shaking) is also characterized by the change in tube characteristics caused by the mechanical action.

The mechanical strength of tubes depends in large measure on the assembly methods used. In manufacturing tubes of increased strength, especial attention should be devoted to the quality of part welding, the treatment of the glass, the strength of connections between the parts, etc.

f) Climatic endurance of receiving-amplifying tubes. The basic climatic factors capable of affecting the operation of tubes are temperature, ambient humidity, and atmospheric pressure.

An increase in ambient temperature increases envelope temperature, resulting in a certain increase in temperature of tube parts (plates, cathodes, etc.). When this happens, there may be an increase in the liberation of gases from the glass and other parts, which in turn may impair the vacuum in the tube and cause it to fail. In additon, where there are sharp temperature variations (thermal shock), the glass of the envelope and mount may be fractured.

A considerable increase in ambient humidity may cause the appearance of leakage currents on the outer surface of mounts and bases. In addition, where tubes and bases remain for extended periods of time under high-humidity conditions, the strength of the bond between the base and tube envelope may be destroyed.

It is also necessary to allow for the possibility of operation of tubes under conditions of sharply decreased atmospheric pressure (for example, at altitudes high above sea level). At lowered pressure, cooling of envelopes owing to air convection is impaired, which is equivalent to operating the tubes at an elevated temperature. In addition, under considerably lower pressures, there is a possibility of

discharges between exposed portions of tube leads to which high voltages are applied.

The climatic resistance of tubes depends basically upon the techniques used in manufacturing them.

g) Tube lifetime. The lifetime or useful service life of tubes is the name given to the period of time (normally measured in hours) of tube operation during which the basic characteristics for the given type of tube do not change by more than a predetermined value, so that normal tube operation in circuits is provided for.

These basic parameters which may be used to judge the serviceability of tubes are: for diodes — the internal resistance; for voltage amplifiers — the transconductance; for output tubes operating in power-amplification circuits — the output power; for frequency-conversion tubes — the conversion transconductance. In addition, for many types of receiving-amplifying tubes, the inverse grid current may be used as a quantity determining the serviceability of tubes.

The necessary tube life is determined by the function of the devices in which the tubes are utilized. For the majority of common types of receiving-amplifying tubes, the guaranteed tube life is 500-1000 hr. Certain types of tubes, operating in so-called one-shot devices, have lives measured in dozens of hours. On the other hand, much radio equipment demands tubes with lifetimes on the order of 5000 - 10,000 hr. (long-distance telephony, radio-relay links, while amplifiers installed in submarine cables require tube lives of up to 100,000 hr). The factors causing tubes to fail as a result of extended operation may be classified into two basic groups. The first group includes mechanical damage to tube parts. Here we have burning out of cathodes and heaters, failure of bonds between tube parts and leads, short circuiting of electrodes caused by electrode deformation, or by foreign

conducting particles getting between electrodes (for example, upon crumbling of the oxide layer of the cathode or getter), cracks in the envelope or mount of a tube.

The second group includes changes in tube electrical properties leading to a deterioration in characteristics. The majority of tube electrical characteristics are connected with cathode emission properties. During operation of tubes, cathode emission properties deteriorate as a result of several physical processes affecting the concentration of free barium in the oxide layer. Such processes may be evaporation of the barium; oxidation of barium by gases liberated during tube operation as a result of decomposition of oxides and other chemical compounds present on the surfaces of tube parts; the effect of secondary emission from tube insulators; processes in the oxide layer and on the base layer that may lead to the formation of a barrier layer on the base-layer — oxide interface.

The second factor causing a deterioration of tube characteristics is the change in contact potential between the control grid and cathode of the tube, occurring as a result of a change in the work function of the grid surface. The change in contact potential may take place so as to produce either an increase or a decrease. If during operation of a tube, the grid surface is activated owing to deposition of active materils from the cathode, the work function will decrease, and the plate-grid characteristic curve will be shifted to the left. At the same time, the converse process may occur — deactivation — i.e., an increase in grid work function owing to poisoning of the active layer on the grid by the oxygen evolving from hot tube parts, and owing to evaporation of this active layer from the grid. As a result, the contact potential may change over a 2-3 v range, and the tube characteristic curve may be shifted within the same limits.

Factors leading to tube failure include a gradual deterioration

in the resistance of interelectrode insulation owing to the deposition of conducting films upon insulators, and also a drop in the insulating qualities of the alundum coating of the heater, occurring as a result of physical-chemical processes in the alundum, and on the heater base-layer — alundum interface. Deposition of active materials upon the grid may lead to the appearance of inverse grid current owing to thermionic emission.

The factors enumerated lead, as a rule, to a slow, gradual change in tube characteristics. In addition, tubes may fail owing to accidental factors associated in most cases with defects in manufacturing processes for individual tubes. Such tubes normally fail in the first few dozen hours of operation.

Since tube life depends upon a large number of different factors, the effect of which may not be the same for all specimens of a single type of tube, in practice, the life of individual tubes may differ sharply. Thus, the guaranteed tube life refers not to any specific tube of a given type, but to the majority of tubes from a manufactured lot. The number of tubes whose life proves to be not less than the warranted life, expressed as a percentage of the total number of tubes in a given batch, is called the tube reliability, for example, if of 1000 tubes of a given type no fewer than 900 have the predetermined life, the reliability is 90 per cent.

Of the total number of tubes having a predetermined lifetime, a large number may operate beyond the warranted time. An increase in the period of time that tubes are operated corresponds, for a given type of tube, to a decrease in reliability.

High tube reliability is especially important where the tubes are utilized in devices containing a large number of tubes (up to several thousand in computers) and in devices where it is difficult

or impossible to replace tubes that have failed. Thus, tubes are produced for such devices having reliabilities close to 100 per cent (98 and 99.8 per cent). High tube reliability is basically achieved as a result of careful quality control of the starting materials and of the tube-manufacturing process, and by testing these tubes under conditions more rigorous than those under which they are to be used.

10-3. STRUCTURAL FEATURES OF RECEIVING-AMPLIFYING TUBES.

The nature and structural features of any type of receiving-amplifying tubes are determined basically by the following factors: 1) the maximum permissible interelectrode voltages; 2) the working frequencies and band width; 3) the power developed on the electrodes of each tube; 4) the method of connecting the tube into a circuit; 5) the requirements for meahanical strength and stability for the tubes, as well as the climatic conditions under which they are to be used; 6) the specifications for tube life and reliability.

Another problem is the question of the desirability of combining two or more tubes in a single envelope.

Let us consider the basic structural features associated with meeting the tube specifications enumerated.

An increase in interelectrode voltages makes it necessary to take special measures to prevent interelectrode sparking along insulation, and for decreasing leakage currents. The majority of receiving-amplifying tube types operate at interelectrode voltages not exceeding 300-500 v. At such voltages, all electrodes of a tube may be mounted on common mica insulators, and the leads of the electrodes may be placed in a common mount. The electrode leads between which the maximum voltage is applied should be located as far apart as possible within the limits of the mount dimensions.

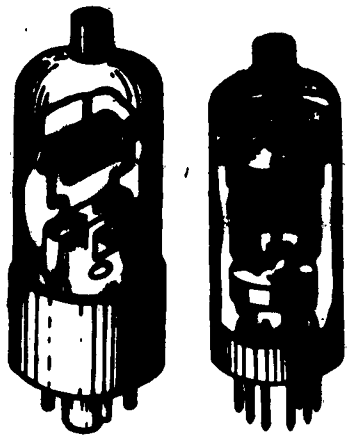

Fig. 10-3. Examples of high-voltage-rectifier construction.

An exception is formed by low-power high-voltage tubes, where the inverse voltage may reach tens of thousands of volts. With such rectifiers, the plate and cathode leads are normally located in separate tube mounts, or the plate lead is soldered directly into a cap on the envelope (Fig. 10-3). The assembly of high-voltage rectifiers on mica insulators common to plate and cathode should be avoided. High-voltage rectifiers are always made in single-plate form, since when the rectifiers are used in a full-wave circuit, large potential differences appear between the plates, which makes it difficult to insulate the plates from each other.

There are substantial structural differences in tubes designed to work in radio frequency amplification and conversion circuits (at frequencies of up to dozens of megacycles. Pentodes are the most common RF amplifier tubes). Such pentodes should have small interelectrode capacitances. The input capacitance of a pentode is the capacitance between the control grid and the electrodes on which, under working conditions, there are no RF varying voltages, i.e., the cathode and the screen and suppressor grids, which in measurement are shorted to-

61-10r

gether. The output capacitance is the capacitance of the plate with respect to the cathode, screen, and suppressor grids, interconnected. The transfer capacitance is measured between plate and control grid with the remaining electrodes grounded.

A considerable decrease in transfer capacitance is provided by the utilization of very fine screen grids which provide strong shielding of the plate from the control grid. The screen-grid fill factor for the majority of RF pentodes is on the order of 0.15-0.20. In addition, in RF pentodes, it is necessary to utilize special shielding of parts associated with the control grid from parts associated with the plate, since the value of transfer capacitance to a large degree depends upon the capacitance between the plate and control-grid leads. In certain cases, the plate and control-grid leads are arranged in this fashion: the plate lead goes through the base mount, and the grid lead through a cap, or _vice versa_.

Internal shielding is accomplished with the aid of shields connected to the suppressor or screen grid. The shape of the shield is so chosen as to shield leads from each other, as well as to shield leads from the electrodes themselves (Fig. 10-4). In the case of single-base tubes, in which the control grid and plate are led out through a single mount, special shielding of the tube mount and base is required.

With good internal shielding, the value of the transfer capacitance is determined basically by the penetrance of the screen and suppressor grids. With sufficiently fine screen grids, the transfer capacitance may be dropped to values on the order of thousandths of a micromicrofarad, and in any case should not exceed 0.03 $\mu\mu f$.

Similar measures are taken to decrease transfer capacitances in RF mixers and converters.

The need for utilizing fine screen grids leads to a decrease in

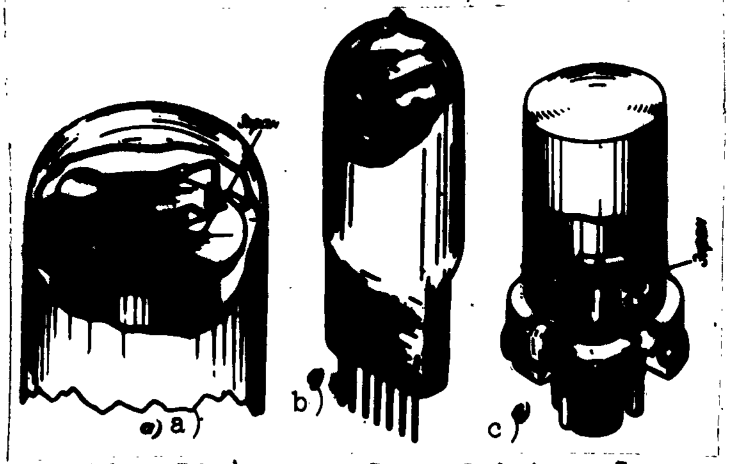

Fig. 10-4. Examples of internal-shield construction.

the current-distribution factor for RF tubes. In most RF pentodes, the current-distribution factor (ratio of the current to screen-grid current) has values ranging from 2.5 to 4.

A special group of RF tubes is formed by tubes used in broadband amplifiers operating with amplified-frequency band widths on the order of several megacycles. Such amplifiers are widely utilized in television, radar, multichannel telephony, etc.

The theory of broadband amplifiers gives the following relationship between the product of the stage gain K and the amplified band width B_{max} and the parameters of the amplifier tube:

$$KB_{\text{макс}} = \frac{S}{2\pi(C_{\text{вх}} + C_{\text{вых}})}. * \qquad (10\text{-}4)$$

This is called the gain-bandwidth product. From this relationship it is clear that the gain-bandwidth product depends solely upon the transconductance of the amplifier tube and the interelectrode capacitances,

* [$K_{\text{макс}}$ – K_{maks} – $K_{maksimal'nyy}$ – $K_{maximum}$
$C_{\text{вх}}$ – C_{vkh} – $C_{vkhodoy}$ – C_{input}
$C_{\text{вых}}$ – C_{vykh} – $C_{vykhodnoy}$ – C_{output}.]

61-102

and is not connected with the resistance of the load on the stage. Evidently, the higher the gain-bandwidth product, the more gain may be obtained at a given bandwidth. Consequently, tubes designed to operate in wideband amplifiers should have the maximum possible ratio of transconductance to the sum of the input and output capacitances.

The second major factor determining the suitability of tubes for utilization in wideband amplification is the magnitude of the tube internal noise.

With a wide band of amplified frequencies, the value of the noise voltage rises. Thus, tubes for wideband amplification should have a relatively low noise level.

Finally, tubes for wideband amplifiers should have a very linear working section of the plate-grid characteristic curve, since nonlinear distortions of the amplified signal will create harmonic components covering the entire band of frequencies amplified.

In order to increase the gain-bandwidth product and decrease the internal-noise level, it is primarily necessary to increase tube transconductance without increasing the size of the electrodes substantially. This may be done by moving the grid closer to the cathode and by creating designs that make better utilization of the cathode surface. Movement of the grid closer to the cathode is associated with a simultaneous decrease in spacing and diameter or the grid turns. A decrease in these quantities is limited by considerations of the manufacturing technology for the grids and for assembly of tubes. Tube production technology available in practice presently makes it possible to obtain a minimum grid-cathode spacing on the order of 30μ with a grid-turns wire diameter on the order of 6-8μ. Improved utilization of the cathode surface is achieved by the utilization of flat (or nearly flat) grid-cathode systems, or of systems with compressed grids.

TABLE 10-1.

1) Обозначение лампы	2) Тип лампы	3) S ма/в	4) $C_{вх}$	5) $C_{вых}$	6) $\frac{S}{C_{вх}+C_{вых}}$	7) $R_{э.ш}$
8) 6Ж4	15) Высокочастотный пентод	9,0	11,0	5,0	0,56	—
9) 6П9	16) Высокочастотный пентод для усиления мощности в широкополосных усилителях	11,7	13,0	7,5	0,57	—
10) 6Ж1Б	17) Сверхминиатюрный высокочастотный пентод	4,8	4,8	3,8	0,56	1 800
11) 6Ж1П	18) Пальчиковый высокочастотный пентод	5,2	4,35	2,45	0,76	1 800
12) 6Ж9П	19) Пальчиковый высокочастотный пентод	17,5	8,0	3,0	1,59	350
13) 6Э5П	20) Пальчиковый высокочастотный тетрод	30,5	15,0	2,55	1,73	350
14) 6С3П	21) Пальчиковый высокочастотный триод	19,5	6,4	1,55	2,45	200

1) Tube designation; 2) type of tube; 3) S, ma/v; 4) C_{in}; 5) C_{out}; 6) $S/(C_{in} + C_{out})$; 7) $R_{e. sh}$; 8) 6Zh4 [6AC7]; 9) 6P9 [6AG7]; 10) 6Zh1B [REGE]; 11) 6Zh1P [6AK5]; 12) 6Zh9P; 13) 6E5P; 14) 6S3P; 15) RF pentode; 16) RF pentode for power amplification in wideband amplifiers; 17) subminiature RF pentode; 18) miniature RF pentode; 19) miniature RF pentode; 20) miniature RF tetrode; 21) miniature RF triode.

It is possible to obtain a linear working section on the plate-grid characteristic curve by shifting the curve slightly to the left, which corresponds to an increase in current and, consequently, of power dissipated at the plate, and in the case of tetrodes and pentodes, on the screen grid.

Table 10-1 gives the basic characteristics of several modern tubes utilized for broadband amplification.

Tubes for low-frequency voltage amplification are considerably simpler in construction; they require no measures for internal shielding. The most common types of tube for audio-frequency voltage amplifi-

cation are medium- and high-mu triodes and audio-frequency pentodes. The advantage lying with pentodes that makes it possible to obtain higher per-stage voltage gain is balanced by the lower nonlinear distortion available when triodes are used.

Pentodes and beam-power tetrodes are more commonly utilized in audio-frequency power-amplification circuits. The basic requirement applicable to such tubes is that they deliver a maximum amount of power with minimum distortion of the amplified voltages.

In order to meet these requirements, audio-frequency tetrodes and pentodes should have a plate-grid characteristic curve shifted to the left by a large amount, a high transconductance, and a rather linear section of the curve in the working range of control-grid voltages. Shifting the curve to the left makes it possible to obtain a high-amplitude AC plate-current component where the amplifier is operated without plate-current cutoff. It is undesirable, however, to obtain a "left" characteristic as a result of increasing the screen-grid voltage, since in this case it is necessary to increase the plate voltage as well. An increase in plate and screen voltages leads to a rise in the power lost on both the plate and the screen, which decreases the efficiency of the amplification circuit and complicates construction of the tube. Thus, beam-power tetrodes and pentodes for audio frequencies have control grids with comparatively high penetrances, on the order of 0.1, which makes it possible to obtain an adequate shift of the plate-grid characteristic to the left with a not-too-high screen voltage, normally close in value to the plate voltage.

A decrease in screen-grid current is important to audio-frequency output tubes. Since the absolute values of currents in these tubes are considerable, a noticeable amount of power is dissipated on the screen. Thus, it is desirable to have the maximum possible current-distribution

factor in these tubes. An increase in the current-distribution factor, in addition to decreasing power losses on the screen grid leads to an increase in transconductance with no change in heating power or electrode voltages. In this respect, beam-power tetrodes have considerable advantages in comparison with pentodes, and thus they are frequently utilized in audio-frequency power-amplification circuits.

In order to improve current distribution, the screen grid of audio-frequency pentodes is normally manufactured with a relatively small fill factor. This makes it possible to obtain a current-distribution factor ranging from 5 to 6.

Since the penetrance of the first and second grids in audio-frequency tubes is relatively large, the total penetrance has a considerable magnitude, as a result of which a variation in plate voltage has a noticeable effect upon the magnitude of the cathode current. Thus, in audio-frequency tubes, the internal resistance proves to be lower than in tubes with fine grids, and has a value ranging from 30 to 100 kohm.

The value of the transconductance for output beam-power tetrodes and pentodes is normally large, in certain tubes amounting to 10-12 ma/v. In output tubes, high transconductance is obtained from a large cathode surface and a high screen voltage (75-100 per cent of U_a), with a relatively high control-grid penetrance.

Depending upon the structural peculiarities of the circuit in which tubes are to be used, the structure for connecting a tube into a circuit is chosen. In cases where rapid changing of tubes by unskilled personnel is specified (broadcast equipment, measurement equipment, etc.), either tubes with standard bases, or with rigid leads having a strictly fixed arrangement and serving as pins fitting into circuit tube sockets are used. In equipment in which it is convenient to solder tube leads to circuit elements, tubes with flexible leads of sufficient length are

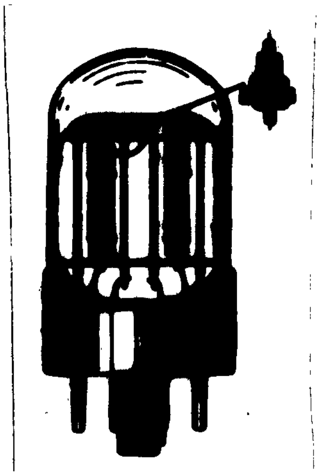

Fig. 10-5. Mounting a dual-diode part with the aid of eyelets.

used. As a rule, flexible leads are used with subminiature receiving-amplifying tubes. The advantages of tubes with flexible leads are reliability of connection into the circuit, structural simplification of circuit application – there is no need for tube sockets. Connection by means of flexible leads somewhat damps the transmission to the tube of various mechanical stresses from the installation.

Strength and stability greatly influence tube design. Vibration is the chief problem. Forces on tube parts during vibration depend upon the frequency of the vibration, upon the amplitude of oscillations, and upon the mass of the parts. In vibration, the direction of the forces is constantly changing, as a result of which the connection between individual tube parts is weakened. This results in an increase in gaps in mounting holes of insulators, between insulators and envelopes, etc. In order to protect tube structures against damage, tube parts must be very tightly assembled. Thus, for example, in tubes designed to operate where considerable vibration is expected, the diameter of grid traverses should exceed by 0.01–0.03 mm the diameter of the corresponding holes in the mica insulator. The same

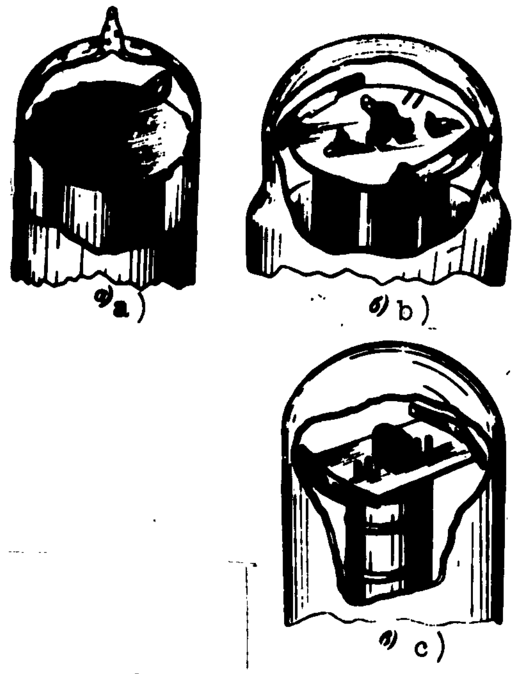

Fig. 10-6. Making tube elements tight in the envelopes.

also applies to the fitting of the indirectly heated cathode into the insulator.

In providing high mechanical strength of tube structures, an essential role is played by the rigidity with which the upper and lower insulators are mounted; all of the parts of the tube are fastened to them. Good results are given where elements are mounted with the aid of eyelets (Fig. 10-5).

When tubes are operated in the presence of heavy vibration, it is very important to prevent tube elements from striking the walls of the envelope, in order to prevent the appearance of oscillations of tube parts at natural (resonant) frequencies. One of the best methods for preventing parts from receiving such blows is tight fitting of the element into the envelope, accomplished either by bending the projecting teeth of the insulators (Fig. 10-6a), or by using special mica side grippers (Fig. 10-6b). Springs are also utilized, firmly fastened

61-107

to the tube element (Fig. 10-6c).

Under working conditions, the possible vibration frequency for tubes almost never exceeds 500-600 cps. Thus, the inherent (resonant) frequencies of individual parts should be considerably higher in order to prevent the appearance of resonant oscillations. For this purpose, the stiffness of part fastenings is increased, and the separation of their supports is decreased.

Increased requirements for tube life and reliability have caused several structural measures to be carried out directed toward weakening the effect of factors leading to tube failure. In order to decrease the effect of the evaporation of active materials from the cathode, it is desirable to decrease cathode temperature within limits providing for normal operation. It is also desirable to decrease heater temperature. As was shown in Chapter Four, a decrease in heater temperature is possible as a result of increasing the fill factor. In addition, it is possible to decrease heater temperature by increasing the absorbing power of the inner surface of the cathode tube by carbonizing it or by an alundum coating. These methods are complicated from the production viewpoint, however, and have not found wide application.

One of the causes of tube failure is burning out of heaters in tubes operating under interrupted-heating conditions, i.e., with frequent connection and disconnection of the heating voltage. Deformation of heaters occurs upon heating and cooling; where recrystalization of the heater wire occurs, this may lead to failure of the heater. The best material for heaters of such tubes is tungsten, which has high strength and a higher temperature of recrystalization than alloys of tungsten with molybdenum.

Of great importance is the proper choice of material for the cathode base layer (in the case where indirectly heated cathodes are used). If the base layer has a large number of reducing additives,

during the first hours of operation a large amount of free barium is reduced, accompanied by rapid evaporation of the barium. Then, as a result of the formation of a barrier layer at the base-layer — oxide interface, owing to the additives, the conditions for barium reduction deteriorate sharply. Thus, the base-layers of indirectly heated cathodes of long-life tubes should utilize nickel with a minimum amount of additives (not more than 0.01 per cent silicon, or not more than 0.03 per cent magnesium).

For stabilization of the grid-cathode contact potential in tubes with indirectly heated cathodes, good results are given by gold-plated grids. In tubes with filaments, magnesium is deposited on the tube parts for the same purpose.

So-called combination tubes, in which two or even three tubes are combined in a single envelope, have found widespread application. The combination of tubes may be organic, in which case the combined tubes have common electrodes, or an independent arrangement may be used. In the first case, the tubes are designed for utilization only in a limited number of types of circuits requiring interconnection of tube electrodes. In the second case, the tubes may be utilized in completely independent circuits.

Twin rectifiers are very common; they are designed for service in full-wave rectification circuits. Two twin-rectifier structures should be distinguished: rectifiers with separate cathodes, and rectifiers with a common cathode and two plates (Fig. 10-7). The majority of dectector diodes are also made in the dual form. The doubling-up of the diodes is explained by the fact that in the majority of modern receivers, in addition to diode detection, there is an automatic gain control (ARCh) circuit, operating in a circuit similar to that of a rectifier, using the second diode. In dual detector diodes, both diodes

of the tube are normally shielded from each other by a shield that has a separate lead through the mount. Sometimes the diodes intended for detection purposes are combined with other tubes working in neighboring stages of the receiver. As an example of such a tube we may take the diode-pentode, which takes the form of a combination of a diode with an RF pentode, and the dual diode-triode, in which there are two detector diodes operating in detection and ARCh circuits, combined with a triode for amplification of the audio voltage. Such tubes, as a rule, have a common cathode.

Dual triodes occupy an important place in the catalog of receiving-amplifying tubes. Twin triodes may be utilized for audio-frequency power amplification, operating in push-pull circuits with low gain, while there are also twin triodes with medium and high gains. Twin triodes find wide utilization in various types of pulse circuits — relaxation oscillators, electronic switches, etc. They may have either common or separate cathodes (Fig. 10-8).

The advantages of combination tubes are first of all savings in weight and size of radio equipment, as well as a certain economy of parts where the parts are common to the combined tube (insulaters, envelopes, mounts, bases, etc.). Where two identical tubes are placed in the same envelope, there is greater uniformity in processing, which in turn decreases differences in characteristics.

Combination tubes, however, are normally complicated to manufacture, and frequently the cost of a combination tube exceeds the cost of two separate tubes.

In constructing combination tubes, it is also necessary to take steps for elimination of interaction. In many cases, forbidden electronic or capacitive coupling between tubes makes it impossible to combine them in a single envelope.

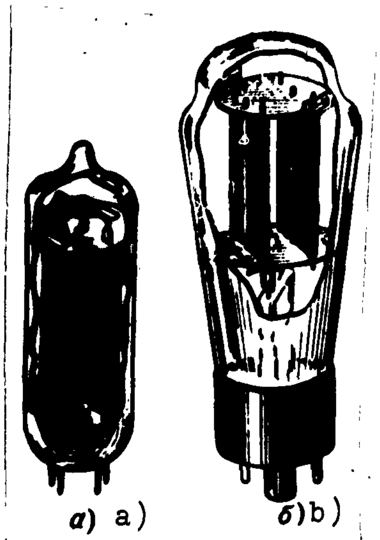

Fig. 10-7. Examples of twin-rectifier construction.
a) With common cathode;
b) with separate cathodes.

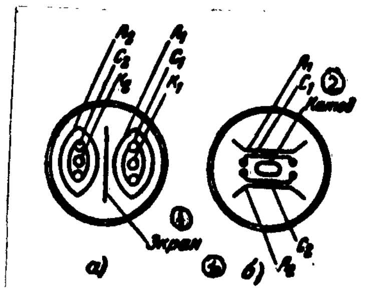

Fig. 10-8. Cross-section of twin-triode construction.
a) With separate cathodes;
b) with common cathode;
1) Screen; 2) cathode.

10.4. CONSTRUCTION OF RECEIVING-AMPLIFYING-TUBE PARTS.

Modern receiving-amplifying-tube designs normally provide for assembly of all interior parts separately from the tube stem; later, the assembled set of parts is connected with the stem. Thus, the structural units of a tube are the set of basic parts, the stem with leads, the envelope, and in certain tubes, the base, serving to provide reliable connection of the tube into the circuit. In many types of tube, there is no base, and the tubes are connected into the circuit directly through the tube lead.

The basis for the assembly of the electrode unit in receiving-amplifying tubes are the mica (or infrequently ceramic) insulators with apertures determining the location of the tube electrodes. The design of parts should provide for the possibility of firm fastening with the insulators and strong connection with leads.

In addition to the units mentioned, tubes also have several auxiliary parts including various types of intratube connections,

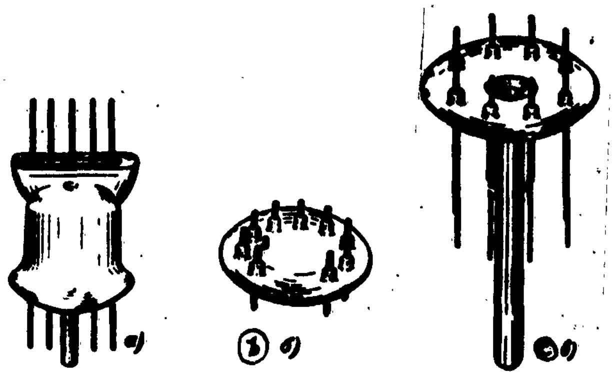

Fig. 10-9. Stem construction for receiving-amplifying tubes.
a) Pressed stem; b) and c) button stems.

supports (or "shelves") for getters, shields, etc.

Below we discuss the structure of several receiving-amplifying tube parts.

Stems for receiving-amplifying tubes. On the stems of receiving-amplifying tubes are fastened the tube mounts, and the electrodes of the tube are linked to the circuit with the aid of leads sealed into the stem (in some cases, a portion of the leads is sealed directly

Fig. 10-10. Sealing-in of leads in subminiature tubes not having stems.

into the tube envelope). On the basis of structure, stems may be classified as pressed and button-type (Fig. 10-9). Press stems are long, which

increases tube dimensions. In addition, long lead length makes for high lead inductance and interlead capacitance. Thus, press stems are found only in older types of glass receiving-amplifying tubes with bases. The advantages of button-type stems are small dimensions, short leads, the possibility of large lead separation and the utilization of leads as pins connecting the tube into a socket (where tubes without bases are used).

Many designs of subminiature tubes have no stem at all, but the leads of these tubes are sealed into the lower portion of the envelope (Fig. 10-10). Leads for these tubes are manufactured of platinum; their external section is tinned. The tube leads are connected into the circuit by soldering them to the appropriate circuit elements.

In press stems and in button stems for glass tubes with bases, an exhaust tube through which the tube is evacuated is sealed in. In button stems of glass and miniature tubes without sockets, there is no exhaust tube, and the tube is exhausted through an exhaust tip sealed into the dome of the envelope.

High-melting glasses, as a rule, are used as stem materials, except in the case of large tubes without bases that have tungsten lead-pins. Since a large number of metal leads are sealed into a stem, considerable internal stresses may appear in the glass, and for this reason the glass may crack, and for this reason lead glasses, which are able to withstand heat treatment (for example, ZS-4 glass), are widely utilized in the manufacture of stems for receiving-amplifying tubes. Another advantage of these glasses is low electrical conductivity at high temperatures.

Envelopes for receiving-amplifying tubes. The structure of blanks for envelopes depends upon the outside shape of the tubes. For glass tubes with press stems, the envelope blank is longer than the finished

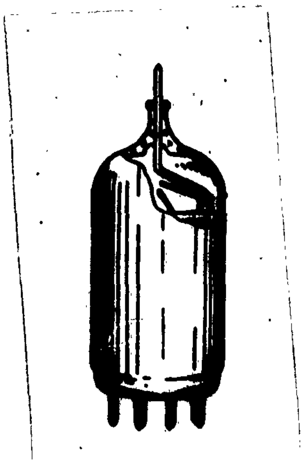

Fig. 10-11. Sealing-in of lead in dome of miniature-tube envelope.

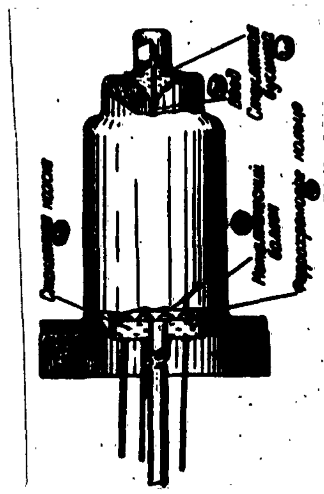

Fig. 10-12. Connection of glass stem and upper lead-in metal tube.
1) Glass stem; 2) lead; 3) glass bead; 4) metal envelope; 5) ferrochrome ring.

envelope, since the lower portion of the blank – the "skirt" – is drawn out after the stem has been sealed into the envelope. The length of envelope blanks for glass tubes with button stems is about the same as the length of the envelope of the finished tube.

In order to bring leads from one or two electrodes out through the top of the envelope, one or two short tips are sealed into it. In envelopes of miniature and subminiature tubes, which are evacuated

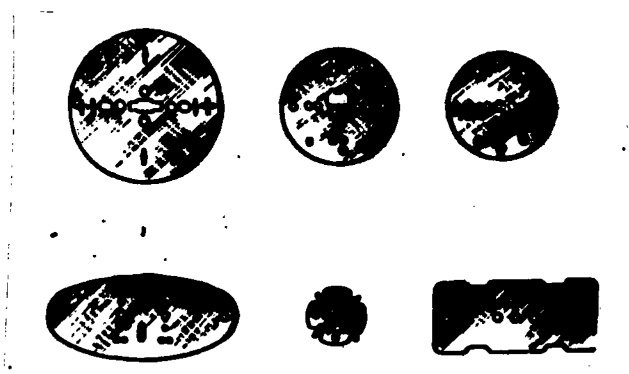

Fig. 10-13. Examples of mica-insulator designs.

through an exhaust tip in the dome, an upper lead is sealed directly into this tip (Fig. 10-11).

In manufacturing envelopes for receiving-amplifying tubes, dolomite glass is used for tubes with press stems and lead glass for miniature and subminiature tubes.

Envelopes for metal tubes are made, as a rule, from low-carbon steel (steel 1010), coated with a special lacquer for protection against corrosion. The envelope is joined with the stem of the tube through a ferrochrome ring (Fig. 10-12). Where there is an upper lead, it is sealed into the top of the metal envelope through a glass bead.

There are several advantages to a metal envelope, basically determined by the fact that the envelope is at the same time an external shield for the tube, and is stronger; these do not compensate, however, for the substantial drawbacks in metal tubes from a production point of view, associated with the impossibility of using high-frequency heating of internal parts, which impairs the quality of the tube and makes it necessary to carry out special treatment of the parts before assembly. Thus, the majority of modern tubes use glass envelopes.

The heat given off by the hot internal parts of a tube heats up the glass of the envelope. The maximum permissible temperature of glass envelopes should not exceed 120-180°C, which corresponds to a heat load on glass envelopes of 0.25-0.35 watt/cm^2.

<u>Insulators for receiving-amplifying tubes</u>. In the majority of modern types of receiving-amplifying tubes, precise mutual positioning of parts is provided by mounting them on special internal insulators (Fig. 10-13). Mica serves as the basic material for insulators (muscovite [common] mica is used); discs of appropriate configuration are stamped out from the material. The number and shape of openings in the discs depend upon the design of tube parts.

Grids, plates, and screens for tubes are fastened into the insulators either with the aid of traverses inserted in circular holes in the discs, or with the aid of tabs bent over after the element has been connected with the insulator (Fig. 10-14). In quite infrequent cases, the plates are fastened to special brackets; this method is seldom used owing to the fact that it does not provide sufficient precision in positioning of the plates.

Depending upon the dimensions of the insulator, the thickness of the mica discs ranges from 0.2-0.5 mm.

Mounting of cathodes. The method used to mount the cathode should position it accurately with respect to the other electrodes, and provide the least possible loss of heat from the cathode owing to cooling through the supports and insulators, and should also insure that there are no lateral deformations (sagging or bending of the cathode as a result of thermal expansion of the cathode.

Filament-type cathodes are positioned accurately with the aid of the methods illustrated in Fig. 10-15. In the first case (Fig. 10-15a),

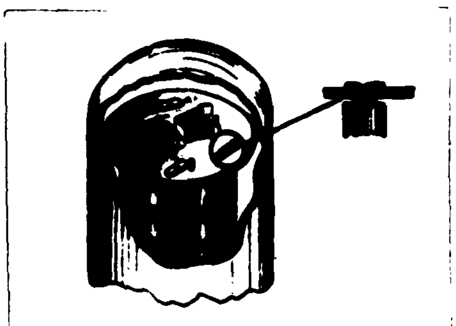

Fig. 10-14. Fastening the plates of a twin diode to the insulator with the aid of taps.

the filament wire rests against the straight side of a hole in the insulator located on the axis passing through the centers of the holes for the grid traverses; in the second case, the wire is set into an

61-107

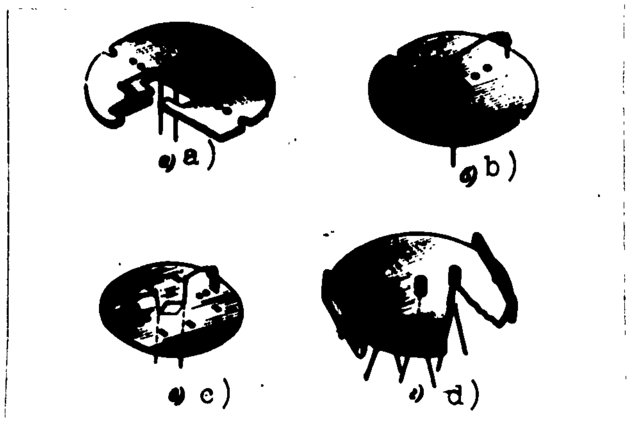

Fig. 10-15. Methods for fastening filaments.

angle in a hole in the insulator (Fig. 10-15b,c). The third method (Fig. 10-15d), in which the position of the cathode is determined by the position of hooks connected to tension springs, is less accurate and is utilized basically in rectifiers.

In order to prevent cathode wire from sagging owing to thermal expansion, tension is applied to it with the aid of springs (Fig. 10-16). Springs are sometimes also utilized for bringing the current into the cathode. In this case, the spring is heated not only by the heat evolving from the cathode, but also by the current passing through the spring itself. Since, at elevated temperatures, spring elasticity decreases sharply, considerable heating of the spring cannot be permitted. Thus, the diameter of the wire used in a spring must be considerably larger than the diameter of the cathode wire. With thin cathodes having a core diameter of up to 30μ, the spring wire will have a diameter 6-10 times the core diameter, while with thinner cores, the spring wire

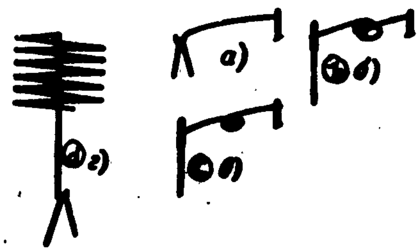
Fig. 10-16. Designs of springs for applying tension to filaments.

61-102

diameter will be 3-6 times the core diameter.

Lever-type springs (Fig. 10-16a) have the greatest rigidity. In order to increase the elasticity of a spring when the wire diameter gets to be fairly large, a single turn or several turns are made in the central portion of the spring (Fig. 10-16b,c); the larger the turns, the less the tensile force, all other conditions remaining the same. Tension springs are fastened either to special hooks (supports), mounted on the mica insulator (Fig. 10-15a), or on the traverses passing through the holes in the insulator (Fig. 10-15b). Spiral tension springs (Fig. 10-15d) rest freely on the insulator.

Springs are manufactured from tungsten, nichrome, and molybdenum wire; it is preferable to utilize tungsten, since at elevated temperatures, the elastic properties of tungsten decline to a lower degree than do those of molybdenum.

Opposite ends of a cathode are rigidly fastened. In tubes with press stems, they are normally fastened directly to supports sealed into the stem. In filament-type tubes with button stems and in subminiature tubes, the insulators are provided with special supports to which the cathode is attached. In tubes with indirectly heated cathodes, the ends of the heaters are fastened to such supports.

Methods for mounting indirectly heated cathodes themselves have been discussed in Chapter Four.

Grids for receiving-amplifying tubes. Spiral-wound grids using two traverses are the most frequently used type; the cross-sectional shape of these grids is determined by the tube design. According to the method used for fastening the wire of the turns to the traverses, which are also classified as welded (Fig. 10-17), where the turns are fastened to the traverses with the aid of electric welding, or as rolled (Fig. 10-18), in which the turns wire is embedded in notches

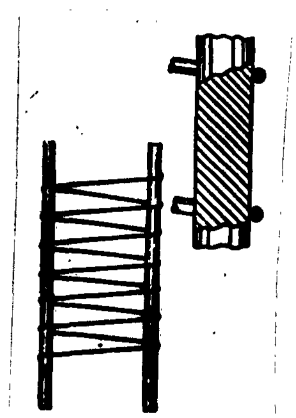

Fig. 10-17. Welded grid.

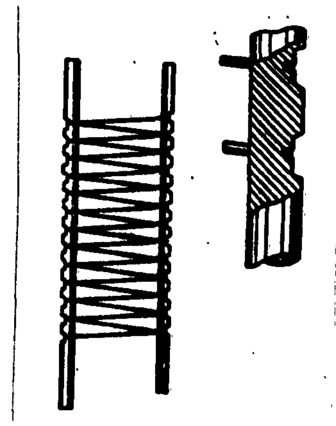

Fig. 10-18. Grid with turns embedded in traverses.

previously cut in the traverses, and then rolled into them. The latter method for manufacturing grids is very accurate and simple from a production point of view. Thus, embedded-type grids are encountered considerably more frequently.

In both cases, the turns are fastened to the exterior surface of the traverse, and the majority of the traverse diameter lies within the grid. This feature makes it impossible to make tubes having cylindrical electrode systems and small grid-cathode spacing. The majority of grid designs have traverse-axis separations exceeding the span (the smallest outside dimension) by a factor of two or more. Fig. 10-19 shows grid cross-section.

Nickel and copper normally are used as materials for grid traverses. Nickel with 2.5 per cent of added manganese is sometimes utilized for thin traverses; this increases the rigidity of the traverses. Copper is utilized for traverses of grids in tubes producing high specific loads on the electrodes, since the high heat conductivity of copper makes for the best cooling. In order to improve weld quality where traverses are electrically welded to other parts of tubes, copper traverses may first be nickel-plated.

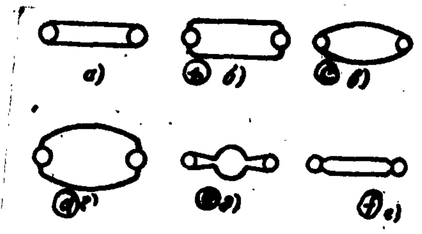

Fig. 10-19. Grid cross-sections.
a) and b) Flat grids;
c) and d) oval grids;
e) and f) shaped grids.

Wire made of tungsten, molybdenum, manganese nickel, and alloys of nickel with tungsten (NIVO), or with molybdenum (NIMO) are used for the grid turns of receiving-amplifying tubes. One of the most common materials in recent years has been molybdenum; its advantage lies in its relatively high rigidity with adequate plasticity, which permits relatively easy fabrication of grids of the required cross-sectional shape. Molybdenum, however, has several faults, of which the basic are the high volatility of molybdenum oxide which tends to poison oxide-coated cathodes. Thus, where molybdenum wire is used to wind grids for tubes having oxide-coated cathodes, the wire is given a protective coating (nickel, gold, silver) to protect the surface against oxidation. The minimum permissible diameter of molybdenum wire in winding grids is 28-30μ. In smaller diameters, molybdenum wire turns out not to be strong enough for manufacturing grids.

Manganese nickel (with 5 per cent manganese) is a good material for grids; it has very high plasticity, and good vacuum properties. Owing to low rigidity, however, it is utilized in grids having a winding diameter of not less than 50μ with a not-too-long turn length.

In modern receiving-amplifying tubes, tungsten, which has high strength and very good vacuum properties, is widely utilized for winding grids, especially those with small winding diameter (50μ and below).

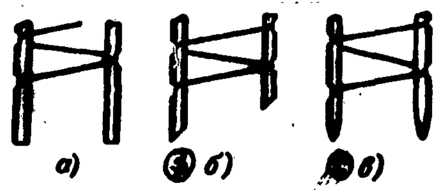

Fig. 10-20. Shapes of free ends of grid traverses. a) Straight cutoff; b) oblique cutoff; c) cigar shape.

Still better properties are possessed by alloys of tungsten with rhenium, which differ from tungsten in that they retain high strength and have high melting properties, but also have high plasticity. This simplifies the fabrication of grids and makes it possible to make grids having any required cross-sectional shape. Owing to the low plasticity of tungsten, it is very difficult to fabricate tungsten grids, and tungsten is generally used in the manufacture of grids with oval cross-sections.

In order to fasten grids into tube insulators, the ends of the grid traverses have no turns at either end. Where a tight fit of traverse into the insulator hole is required, the free ends of the traverse may either be cut off obliquely, or cigar shaped (Fig. 10-20).

Where the diameter of the turns wire is not less than 20μ (with tungsten wire) it is possible to manufacture spiral grids on traverses with embedded turns. With a further decrease in turns-wire diameter, the structural rigidity of the grids, and their ability to hold their shape, turns out to be completely inadequate. Thus, where considerably thinner turns wire is to be used for grids, radically new grid structures are needed. The basic grid structure for grids with thin turns wire (5-8μ) is a rigid frame upon which the turns are placed under tension. This construction may be two-sided, similar to a standard grid, in which case it takes the form of a frame made of two normally

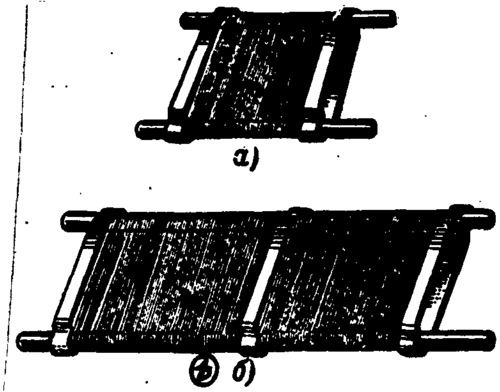

Fig. 10-21. Construction of double sided frame-type grids with two (a) and three (b) pairs of bands.

circular traverses, joined by two (or three) pairs of bands (Fig. 10-21). The bands are either welded or soldered to the traverses, with gold, covering the traverses and bands, serving as the solder. A thin gold-plated tungsten wire is wound on this frame under tension; the ends of the wire are fastened to the traverses. When the system is heat treated, the position of the turns becomes fixed upon the traverses, owing to the mutual diffusion of gold from the surfaces of the turns and the traverses. The diameter of the traverses should be sufficiently large to provide structural rigidity, since the force set up by the tension on all of the turns may cause bending of the traverses. Thus, where long frames are involved, a third pair of bands is used in the center.

A second frame-type grid design is the single-sided grid wound on lat-stammed frames (Fig. 10-22). In such grids, the turns wire and frame are also gold-plated and the turns are joined to the frames by gold soldering. The stamped-frame grid design can provide quite high accuracy in turns spacing.

Materials with good mechanical properties (molybdenum and tungsten) are used in both frame-type grid designs for manufacturing the grids, and thus the necessary structural rigidity is provided.

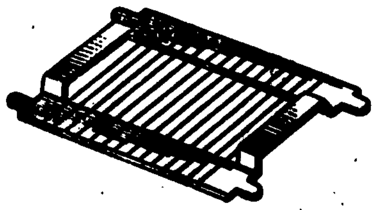

Fig. 10-22. Example of construction of single-sided grid on a stamped frame.

Plates for receiving-amplifying tubes. In choosing plate designs, in addition to the dimensions determining the basic tube characteristics, it is necessary to allow for several additional factors determined by these specific features of the type of tube. These factors are the thermal operating regime of the plate, the method used to mount it into the tube, the requirements for mechanical stability and lifetime of the tube, etc.

For the majority of receiving-amplifying tube designs, the plate is the support for the insulators upon which the remaining parts of the tube are mounted. Thus, the fastening of the plate through the insulators is of essential importance in establishing the required accuracy of positioning of the plate relative to the other electrodes. When the plate is mounted on traverses, positioning is more accurate than when tabs are used for mounting, but tab-mounting does not require the utilization of additional parts (eyelets, locking devices, etc.) to increase the strength of the plate-insulator joint. In many cases, combination mounting using traverses and tabs is used. Where plates are mounted on tabs, the dimensions of the holes in the insulator normally permit the plate to be shifted in the plane of the tabs. Thus, the tabs should be so arranged that such a shift has no effect upon the interelectrode distances determining the tube characteristics.

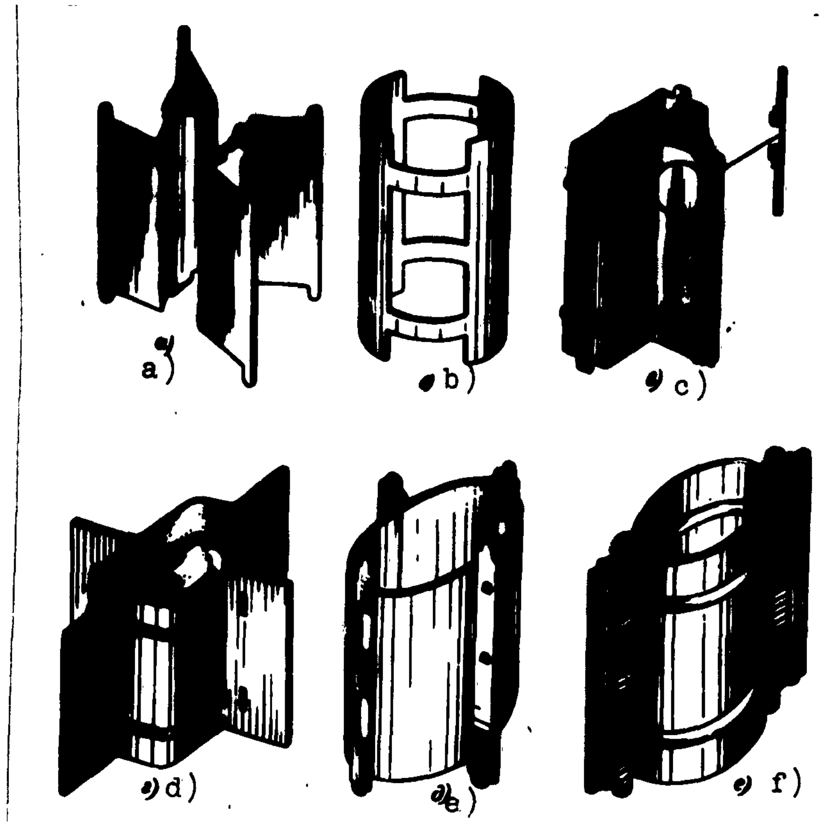

Fig. 10-23. Examples of plate structures.

Materials for the plates of receiving-amplifying tubes are nickel, as well as several laminated materials with coatings on one or two sides such as, for example, aluminum-plated steel, aluminum-plated nickel, nickel-plated steel, etc.

Common (white) nickel is utilized as a plate material in tubes having small specific plate dissipation. For tubes with large specific plate power, carbonized or ground nickel or aluminum-plated steel and aluminum-plated nickel are used; the surface of these materials is carbonized after hydrogen has been driven off; they have a high emission factor. The advantage of aluminum-plated steel is that it is relatively easily outgassed, and liberates little gas during operation of the tube. Owing to large rigidity of this material, however, there are production problems in manufacturing plates from it.

Open plate designs (Fig. 10-23a,b), used to decrease tube output

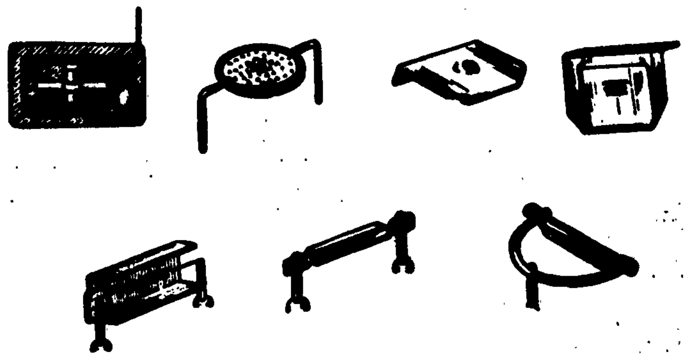

Fig. 10-24. Examples of getter structures.

capacitance, also improve cooling of grids in tubes with high heat loads, and make it easier to control the quality of tube-element assembly. Drawbacks to these plates are their relatively low rigidity and difficulties involved in heating them with high-frequency currents during evacuation.

Closed-design plates are either folded in one piece from a blank (Fig. 10-23c) or made from several stamped sections (Fig. 10-23d, e, f). The individual plate sections are joined together by either spot welding or mechanical stapling (Fig. 10-23c). The advantage of closed plate designs is their rigidity, sometimes increased by means of special ribs. For purposes of quality control during assembly of tube elements, and to improve cooling of grids, plates are sometimes manufactured with windows in side portions, and with a notch in the ends of the working section.

Auxiliary parts. Auxiliary parts involved in the construction of receiving-amplifying tubes are getters, shields, various connections and bases.

In order to obtain a high vacuum in tubes, a mirror-like substance is applied to the interior walls of the envelopes, a substance that actively absorbs gases – a getter. Among such substances are barium, magnesium, and various alloys of these metals with aluminum, thorium,

etc.

A common method for applying the mirror-finish getter to receiving-amplifying tubes is evaporation of the getter material which is used in the form of pressed tablets or spread on special supports — shelves. These "shelves" should be constructed so as to direct the evaporating getter material only toward the free portion of the interior surface of

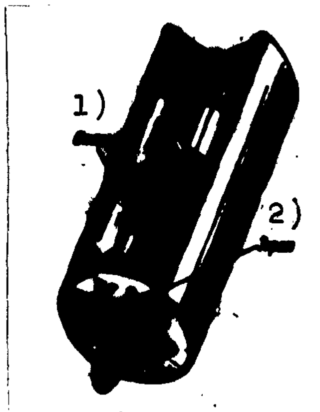

Fig. 10-25. All-glass tube in metal shell with seal-in base plug.
1) Envelope; 2) shell.

the envelope (Fig. 10-24). In order to facilitate heating up of the getter carriers where their surface is small (for example, with a ribbon-type getter) by using high-frequency currents, it is desirable to mount them so that the bearers and the tube elements upon which they are mounted form a closed loop. The size of the getter mirror-surface is established depending upon the volume of the tube, the dimensions of the elements heating up during tube operation, and their working temperatures.

Bases are used in many tube designs to connect the tubes into their circuits. The base takes the form of a cylinder closed at one end by an insulating plastic bottom through which are pressed metal pins connected to the tube leads. The cylindrical base shell is either pressed as a whole from plastic, or assembled from a plastic bottom

61-102

(insert) and a metal body. In the center of the bottom of the base, there is an aligning plug, within which is located the sealed end of the exhaust tip. On the outer surface of the plug there is a key, for which a corresponding notch is made in the center hole of the tube socket, so that the tube may be plugged into the socket in a single position only.

Bases are fastened to the envelopes of glass tubes with the aid of tube-base compounds, while in metal tubes the base insert is fastened to the lower portion of the shell.

Where there is an electrode lead through the top of a tube, a special cap is used for purposes of connection into the circuit; the cap is attached to the envelope with a tube-base compound.

In glass tubes with an external shield, the leads themselves are used as the pins connecting the tube into the circuit. In order to protect the exhaust tip, and to insure that the tube is properly inserted into the socket, a metal base with openings for the lead-pins, and with an aligning plug, is placed on the tube. This base is joined with the cylinder formed by the external metal shell in which the entire tube is located (Fig. 10-25).

10-5. FUNDAMENTALS OF RECEIVING-AMPLIFYING-TUBE DESIGN.

The initial data for designing tubes are the static characteristics, the magnitude of the plate power dissipation, and the cathode heating parameters. In cases where the parameters given are associated with the circuits in which the tubes are to be used, it is necessary to convert them to the data enumerated. In addition, it is necessary to determine the operating regime in which the tubes will have the required parameters.

The design problem amounts to the determination of the geometric

dimensions of all tube electrodes. It is clear that choice of the tube arrangement must precede calculation. For precise tube design, and especially for tubes with grids, it is necessary to take into account many phenomena that determine the passage of current through tubes — the initial electron velocities, contact potential, deviation (deflection) of electron trajectories, etc. Below we give approximate methods for designing tubes of very simple structure.

Rectifier design. The basic data for designing rectifiers normally are the cathode heating voltage, the magnitude of the rectified current and rectified voltage at the given transformer-secondary voltage in the rectifier circuit. In designing tubes, it is necessary to know the value of the peak inverse voltage.

According to the given values of rectified voltage and peak transformer voltage, the peak plate current and plate voltage are found in accordance with formulas (5-37), (5-38), (5-39) and the chart of $n = f(\eta)$ (Fig. 5-22). From the peak-plate current value that is found and the given rectified-current value, the cathode is designed according to the circuit shown in Chapter Four.

Then, substituting the values of peak plate current and peak plate voltage into the formula expressing the three-halves law, we find the plate-cathode separation for a planar system, and the plate radius for a cylindrical electrode system.

For a rectifier having a loop-type filament and a planar plate the perveance G in the three-halves law will involve a value of effective plate surface equal to the area of the projection of the cathode upon the plate, where the width of the projection equals twice the plate-cathode spacing. For the case in which $a \gg 2r_a$

$$F_{a.э\phi} = \frac{4lr_a}{a}[2na - (2n-1)r_a],* \quad (10\text{-}5)$$

where a is the distance between cathode mounting points; n is the number of loops of the cathode; l is the height of a loop; r_a is the plate-cathode spacing.

Then G has the following value:

$$G = 9{,}32 \cdot 10^{-6} \frac{l[2na - (2n-1)r_a]}{ar_a}. \quad (10\text{-}6)$$

for the case in which $a < 2r_a$, the effective plate surface equals:

$$F_{a.э\phi} = l[(2n-1)a + 4r_a] \quad (10\text{-}7)$$

and

$$G = 2{,}33 \cdot 10^{-6} \frac{l[(2n-1)a + 4r_a]}{r_a^2}. \quad (10\text{-}8)$$

Substituting the calculated value of G into the three-halves law we find from it the quantity r_a. Since the ratio of r_a and a is not specified in the computations, it is necessary to solve two equations: one for the case in which $a > 2r_a$, and one for the case in which $a < 2r_a$. It is necessary to pick that one of the two solutions that satisfies the initial conditions.

Since alternating current is normally used to heat rectifier cathodes, there is no need to allow for nonunipotential conditions in the computations.

For a cylindrical system with an indirectly heated cathode, geometric design of the plate proceeds according to formula (5-20). The effective plate surface in this case equals $2\pi r_a l_k$, or l_k equals the length of the oxide coating on the cathode. Substituting this

* [$F_{a.e\phi}$ - $F_{a.ef}$ - $F_{\text{anod effektivnyy}}$ - $F_{\text{effective plate}}$.]

value of $F_{a.ef}$ into Eq. (5-20), we obtain

$$I_a = 2{,}33 \cdot 10^{-6} \frac{2\pi l_k}{r_a \beta^2} U_a^{\frac{3}{2}}, \tag{10-9}$$

whence

$$r_a \beta^2 = 14{,}65 \cdot 10^{-6} l_k \frac{U_a^{\frac{3}{2}}}{I_a}. \tag{10-10}$$

The computed value of the product $r_a \beta^2$ is divided by r_k, and Table 5-1 is used to determine the ratio r_a/r_k and, consequently, the plate radius as well.

High-voltage rectifiers designed for inverse voltages over 10 kv frequently use cylindrical plates and directly heated cathodes, made in the form of parallel filaments or spirals installed coaxially with the plate (Fig. 10-3). Plate-geometry computations for a spiral-cathode — cylindrical-plate system may be carried out as for a cylindrical system with an indirectly heated cathode, introducing a correction that takes into account the ratio of the spiral pitch to the plate-cathode spacing. Thus, the equation for the plate current will take the following form:

$$I_a = 2{,}33 \cdot 10^{-6} \frac{F_{a.эф}}{r_a^2 \beta^2} f_0 U_a^{\frac{3}{2}}. \tag{10-11}$$

The quantity f_0 is a function of the ratio $t/(r_g - r_k)$, where t is the spiral pitch. Figure 10-26 shows a graph of this function.

Triode design. The initial data for triode design are normally the heating voltage (in some cases, the heating current is given as well), the amplification factor and transconductance under the given operating conditions, determined by the plate and grid voltages.

With the system of electrodes chosen, it is necessary to perform the calculations for the cathode. With the heating voltage and current given, the cathode is designed for the given electrode system according

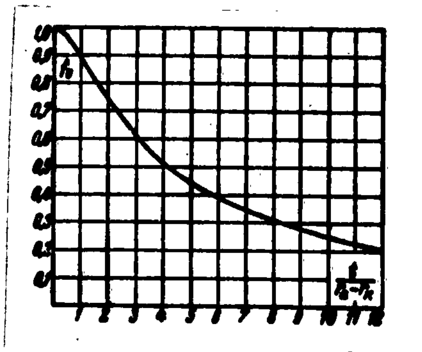

Fig. 10-26. Graph of the function $f_0(t/r_a - r_k)$.

to the known power. Here the diameter of the cathode tube is found, and the length of the oxide coating is determined. If the heating power is not given, the surface area of the oxide-coated portion of the cathode is found either on the basis of the cathode working current, or according to the given transconductance.

In the first case, the cathode working current, equal to the rated plate current, is found from the expression

$$I_a = \frac{2}{3} S U_A \tag{10-12}$$

and the magnitude of the oxide-coated cathode surface is found as

$$F_{\text{окс}} = \frac{I_a}{I_p}. \quad *$$

In designing a cathode, it is possible to use the formulas for transconductance to find the grid-cathode spacing for a planar system, or the grid radius for a cylindrical system. For a cylindrical-cathode — oval-grid system, the transconductance may be computed on the basis of formula (6-46), allowing for the cathode coverage factor γ:

* [F_{oks} - F_{oks} - $F_{oksidirovannyy}$ - $F_{oxide\text{-}coated}$.]

$$S_a = 2{,}33 \cdot 10^{-6} \gamma \frac{2\pi l_{oxc}}{r_c \beta^2} U_A^{\frac{1}{2}}. \tag{10-13}$$

An oval-cathode — oval-grid system may be treated as a planar system, where the effective plate surface equals the surface of the working sections of the oxide-coated portion of the cathode.

If computations for tube geometry are carried out according to a given value of transconductance, when the cathode-tube diameter has been chosen, it is necessary to get the grid-cathode spacing and to determine the length of the oxide-coated section from the formula for the transconductance.

Triodes using filaments are designed as if they were directly heated diodes. In the computations, it is necessary to allow for a correction accounting for cathode-potential nonuniformity.

The next step in the design is to find the grid parameters. The grid pitch is so chosen that no "island" effect appears. In order to do this, it is necessary to satisfy the condition $t < 2r_{s.k}$. Where possible, the relationship between the pitch and the spacing of grid and cathode should be so chosen that the grid pitch will be less than this spacing.

In order to determine the dimensions of the plate, it is necessary to have the grid fill factor α. For low-mu triodes (up to 10) the value of α is selected in the 0.05 to 0.15 range, while for medium- and high-mu triodes, it ranges from 0.2 to 0.3. With the grid fill factor available, it is possible to use the formulas for amplification factor to find the plate dimensions. For cylindrical and planar systems, the value of r_a may be calculated directly from formulas (6-47) and (6-48). Where the electrodes are complex in shape, the quantity r_a is best found from Ollendorf's formula (6-50).

In practice, tube element dimensions are frequently precalculated, using as a basis tubes with known dimensions and characteristics, con-

verting these quantities so as to produce tubes with the required parameters. Such a conversion is possible where the tube electrode shapes are preserved. When such a conversion is performed, corrections to tube dimensions are made on the basis of the parameters of sample specimens constructed in accordance with a preliminary calculation.

Correction of element dimensions with a view toward obtaining a predetermined amplification factor may be done by two methods: by changing the grid fill factor, and by changing the plate-grid spacing. Clearly, in order to increase the amplification factor, it is necessary to increase the "density" of the grid, i.e., to decrease its penetrance. A change in grid fill factor is achieved either by changing the pitch, or by changing the diameter of the grid wire. Similarly, the plate-grid spacing may be increased in order to increase the gain, and decreased to decrease the gain.

Over a comparatively large range of amplification-factor variation, it is possible to assume that it is a linear function of the plate-grid spacing, all other electrode dimensions remaining the same. Thus, for conversion purposes, it is possible to utilize the following formula:

$$\frac{\mu''}{\mu'} = \frac{(r_a'' - r_c)}{(r_a' - r_c)}. \tag{10-14}$$

In this formula, μ' and r'_a are the gain and plate dimension of the known tube; μ'' is the required gain, corresponding to the unknown quantity r''_a.

In order to change the gain by means of grid pitch and grid-wire diameter, it is possible to utilize a formula that gives sufficiently accurate results where the changes in amplification factor do not go outside of the 15-20 per cent range:

$$\frac{\mu''}{\mu'} = \frac{d_c''}{d_c'}\left(\frac{t'}{t''}\right)^2. \tag{10-15}$$

In practice, tube parameters are corrected most frequently by changing

grid dimensions, since it is simpler from a production point of view to change the pitch or grid-wire diameter than to change the dimensions of the plate, requiring the manufacture of an expensive tool.

Triode transconductance may be corrected either by changing the grid-cathode separation, or by changing the active length of the cathode. In tubes with oxide-coated indirectly heated cathodes, a change in transconductance by 10-15 per cent may be achieved by changing the length of the oxide coating. In this case, it is necessary to use the fact that the value of the transconductance is directly proportional to the active length of cathode.

Since the transconductance is strongly dependent upon the grid-cathode separation, the normal method used to change it is a change in the spacing. Approximate formulas for making conversions in grid-cathode spacing take the following form:

1. For a planar electrode system

$$\frac{S'}{S''}=\left(\frac{r_c''}{r_c'}\right)^2. \qquad (10\text{-}16)$$

2. For planar electrode systems with a loop-type cathode

$$\frac{S'}{S''}=\frac{r_c''}{r_c'}. \qquad (10\text{-}17)$$

3. For a cylindrical electrode system

$$\frac{S'}{S''}=\frac{r_c''\beta_2^2}{r_c'\beta_1^2}. \qquad (10\text{-}18)$$

In these formulas, the quantities r'_c and β_1^2 refer to the tube with the known parameters, which has a transconductance S', while the quantities r''_c and β_2^2 are the unknowns for the tube with a preassigned transconductance S". When changing the grid-cathode separation, it is necessary to keep in mind the fact that the plate-grid spacing is changed simultaneously and, consequently, the triode amplification factor also

changes. It is possible to neglect this change only where the plate-grid separation considerably exceeds the grid-cathode spacing. Where the grid-cathode spacing is decreased, it is necessary to allow for the possibility of an intensification of the "island" effect.

Triode and pentode design amounts to determining the geometric dimensions of the tube electrodes from the given characteristics. The initial magnitudes involved in the computations are the cathode heating voltage and current, the plate current at the given plate and screen voltages, and control-grid bias, the transconductance at the operating point, and for RF tubes, the values of the interelectrode capacitances. These parameters are now always explicitly given. Thus, for example, the basic parameter of tubes used for power amplification at audio or radio frequencies is the useful power delivered under the given amplifier operating conditions. In such cases, it is first necessary to determine the tube parameter required to provide the given parameters for the circuit in which the tube is to be utilized, by means of an analysis of the operating conditions.

The functional dependence of beam-power tetrode and pentode parameters upon electrode dimensions is expressed by complicated equations: here the same characteristic may correspond to various combinations of electrode dimensions. Thus, in design work for such tubes, it is necessary to supply certain dimensions, which requires a great deal of experience and an acquaintance with available tube designs. Sufficiently accurate calculations for standard tetrodes are in general impractical to carry out, owing to the impossibility of allowing for the effect of secondary electron emission upon the division of currents in the tube.

Tube design begins with the choice of the system of electrodes and the design of the cathode.

In order to compute control-grid geometry, it is necessary to find

the effective potential U_d and to select the current-division factor $\underline{k}$. The value of the effective potential may be determined from the plate-current and transconductance equations for tetrodes and pentodes:

$$I_a = \frac{k}{k+1} G U_{\text{д}}^{\frac{3}{2}},$$

$$S = \frac{3}{2} \frac{k}{k+1} G U_{\text{д}}^{\frac{1}{2}}.$$

The ratio of plate current to transconductance coincides with the expression given previously for the ratio of diode plate current to slope:

$$\frac{I_a}{S} = \frac{2}{3} U_{\text{д}}. \tag{10-19}$$

From this expression, it is possible to determine the effective potential, knowing the plate current and transconductance.

The choice of current-division factor depends upon the given values of plate and screen voltage, transconductance, and transfer capacitance, and is made by analogy with known tube types.

With the current-division factor selected, it is possible to utilize the plate-current equation to find the value of G, which is determined by the dimensions of the control grid and cathode. Next, the grid-cathode spacing is found for a planar system, and the grid radius for a cylindrical system, as in designing a triode.

In order to carry out the computations for the control-grid – screen-grid separation, it is necessary to know the penetrance D_1 of the control grid. For RF tubes, in which the effective potential is determined by the voltages on the control and screen grids, the value of D_1 may be found from the expression for the effective potential

$$U_{\text{д}} = U_{c1} + D_1 U_{c2}. \tag{10-20}$$

whence

$$D_t = \frac{U_k - U_{c1}}{U_{c2}}. \qquad (10\text{-}21)$$

the quantity that is the reciprocal of the control-grid penetrance:

$$\mu_T = \frac{1}{D_t}$$

is called the triode amplification factor. In designing beam-power tetrodes and pentodes for audio frequencies, this quantity is normally supplied, since it determines the shift in the plate-grid characteristic curve.

With the value of triode amplification factor determined, it is necessary to choose the pitch and thickness of the control-grid turn. The value of the pitch is so calculated as to avoid the "island" effect. In RF tubes, in order to decrease the input capacitance, it is desirable to move the screen grid further away from the control grid, and thus the grid fill factor should not be made especially large, and in any case should be not more than 0.2-0.25.

Using the known geometry of the control grid and the triode amplification factor, we find the screen-control grid spacing very simply, and with sufficient accuracy, from Ollendorf's formula (6-50). Rearranging this formula for the case under consideration, it is possible to obtain the following expression for the unknown spacing between the grids:

$$r_{c2} - r_{c1} = \mu_T \cdot t_1 T - \Delta t_1. \qquad (10\text{-}22)$$

In this formula, μ_t is the triode amplification factor of the tube; t_1 is the control-grid pitch; T and Δ are functions of the grid fill factor determined according to the graph in Fig. 6-28.

For beam-power tetrodes, by carrying out this calculation, we

completely specify the dimensions of the screen grid, since the screen-grid pitch equals the control-grid pitch, while the grid-wire diameter will either equal or be slightly smaller than the diameter of the control-grid wire. For pentodes, choice of screen-grid penetrance is associated with two basic factors: providing good shielding of the plate from the control grid in accordance with the predetermined value for the transfer capacitance, and the value of the current-division factor. In addition, the shape of the plate characteristic curve is a function of screen-grid pitch along the section where it changes over from electron-return to direct interception operation. Suppressor-grid geometry in large degree influences these characteristics in pentodes. It is complicated and not very accurate to calculate mathematically the optimum screen and suppressor dimensions and plate-screen spacing. Thus, the geometry of the space between the screen grid and the plate is normally found experimentally for beam tetrodes and pentodes. Here, in the case of RF pentodes, it is exceptionally complicated to decrease the transfer capacitance, which leads to the necessity of choosing as small as possible a value for the grid penetrance. For this purpose, in RF pentodes, the screen-plate spacing should be increased, insofar as acceptable tube dimensions permit this to be done. These steps cause an impairment in the shape of the plate curve; the beginning of the flat section of the characteristics is shifted toward higher plate voltages.

In selecting the geometry of the screen-grid – plate spacing in audio pentodes, especial attention should be paid to increasing the current-division factor, and obtaining a sharper transition point on the plate curve. In order to do this, the screen-grid fill factor is made small, on the order of 0.08–0.12. The screen-suppressor separation and the suppressor-plate separation should be made small. Clearly,

the screen-suppressor spacing may be computed from the formula for the current-division factor, if the pitch and wire diameter have been chosen for both grids. Knowing the plate-suppressor separation, which normally will be somewhat larger than half the grid pitch, it is possible to determine the penetrance of the suppressor and, consequently, the effective potential in the plane of the suppressor grid. From the formula for current division, the value of C_1 is determined:

$$C_1 = \frac{k}{\left(\frac{U_{a3}}{U_{c2}}\right)^{\frac{1}{2}}} \tag{10-23}$$

while r_{c3} is found on the basis of formulas (8-8) and (8-9) for C_1.

In practice, it is frequently necessary to correct the current distribution on the basis of specimen models of the tubes. The simplest method for correcting current distribution is to change the screen-grid fill factor. Where the dimensions of the remaining electrodes remain unchanged, the screen-grid fill factor α_2 required to obtain the necessary current-distribution factor k_2, may be determined according to the following formula:

$$\alpha_2 = \alpha_1 \frac{k_1}{k_2}, \tag{10-24}$$

where α_1 and k_1 are the fill and current-distribution factors for the specimen model under investigation.

Of practical interest is the recalculation of the screen-grid fill factor in order to change the current distribution while maintaining the grid penetrance constant. As we have already shown, upon recalculation of the amplification factor, it is possible to make use of (10-15)

$$\frac{\mu_1}{\mu_2} = \frac{\alpha_2'}{\alpha_2''}\left(\frac{r''}{r'}\right)^2.$$

It follows from this expression that constant amplification factor (or penetrance) should result when the following equation is obtained:

$$\frac{d_c''}{(l'')^2}=\frac{d_c'}{(l')^2}. \tag{10-25}$$

since

$$\frac{d_c}{l}=a,$$

Equation (10-25) may be written in the following form:

$$\frac{a_2}{l''}=\frac{a_1}{l'}. \tag{10-25a}$$

Substituting into this equation the value of α from expression (10-24) and simplifying, we obtain:

$$\frac{k_2}{k_1}=\frac{l'}{l''}. \tag{10-26}$$

It is clear from relationship (10-26) that in order to increase the current-distribution factor while maintaining the screen-grid penetrance constant, it is necessary to decrease the screen-grid pitch, which entails a simultaneous decrease in the diameter of the grid wire.

Plate design. The amount of power dissipated by the plate at its maximum permissible temperature consists of the power carried to the plate by the electrons (the plate current) and the power absorbed by the inner surface of the plate owing to radiation of interior tube elements – cathodes and plates. Under steady-state conditions, the power dissipated by the plate will equal:

$$P_a=U_{\text{а.эфф}}I_{\text{а.эфф}}+q_kP_n+q_cP_c. \tag{10-27}$$

where P_n is the cathode heating power; P_c is the power radiated by the grids; q_k and q_c are factors indicating the relative proportions of power radiated by the cathode and grids respectively that is ab-

sorbed by the plate. The coefficients q_k and q_c are strongly dependent upon tube geometry and the properties of the inner surface of the plate.

In cylindrical structures (or nearly cylindrical structures), where the length of the plate considerably exceeds the interelectrode spacings, where closed plate structures are used, it is possible to assume that $q_k \approx 1$. The same is true with respect to q_c where the conditions prevailing are such that the grids do not have radiators to improve cooling. With open plate structures, and where the length of the plate approximates the plate-cathode spacing, the coefficients q_k and q_c prove to be considerably less than 1. The power radiated by the plate consists of the power radiated by its outer surface, and the power radiated by its inner surface. The relationship between the magnitudes of the powers radiated by the inner surface and the outer surface depends upon the ratio of the length of the plate to its diameter. For a plate length that is five times the diameter (or the span for oval and planar designs), the power radiated by the inner surface does not exceed 10 per cent of the total power radiated by the plate, and it may be neglected.

In order to insure that the plate temperature does not exceed permissible values under steady-state conditions, the plate should . e a radiating surface that is adequate to dissipate all of the incoming power at this temperature. The size of this surface is determined in terms of the ratio between all power radiated to the power radiated per cm^2 of plate material at the given temperature:

$$F_{изл} = \frac{P_a}{P_{уд}}.* \qquad (10\text{-}28)$$

* [$F_{изл}$ - F_{izl} - $F_{izlucheniye}$ - $F_{radiation}$.]

TABLE 10-2

1) Материал анода	2) Допустимая температура, °C	Допустимая удельная мощность $P_{уд}$, *вт/см²* 3)
Белый никель	700—850	1,0—1,5
Белый никель в лампах с оксидным катодом	400—450	0,2—0,3
Черненый никель	700—850	2,8—4,2
Черненый никель в лампах с оксидным катодом	400—450	0,8—1,2
Молибден белый	1 100—1 150	4,0—6,0
Молибден цирконированный	900—950	8,0—10,0
Тантал	1 250—1 300	8,0—8,5
Графит	1 400—1 500	45,0—[illegible]

1) Plate material; 2) permissible temperature, °C; 3) permissible specific power P_{ud}, watt/cm³; 4) white nickel; 5) white nickel in tubes with oxide-coated cathode; 6) carbonized nickel; 7) carbonized nickel in tubes with oxide-coated cathode; 8) molybdenum, white; 9) zirconium-coated molybdenum; 10) tantalum; 11) graphite.

The values of P_{ud} for various materials are shown together with the permissible temperatures in Table 10-2.

The values given in the table for the permissible specific power are correct for an ambient temperature of 20°C.

Formula (10-28) is used to compute the required radiation surface; it is also possible to calculate the surface area of the plate cooling fins.

$$F_p = F_{изл} - F_a.* \qquad (10\text{-}29)$$

In plate design, it should be assumed that both sides of a fin are radiators.

---

* [F_p - F_r - F_{rebro} - F_{fin}.]

Envelope design. All of the power liberated in a tube, with the exception of the heat carried off by the leads, is applied to the envelope. In turn, the envelope transmits this power into the surrounding space, first, owing to convective cooling, and, second, owing to radiation. In the majority of tube designs, the power carried off by the leads may be neglected in approximate calculations pertaining to envelopes. The connection between the power given up by an envelope and its temperature under normal atmospheric-pressure conditions is established by the equation

$$P_б = A(T_б - T_0)^{\frac{5}{4}} + F_б \varepsilon \sigma (T_б^4 - T_0^4).* \qquad (10\text{-}30)$$

In this equation, T_0 is the ambient temperature, T_b is the mean envelope temperature, A is a coefficient determined by envelope geometry; F_b is the total envelope surface; ε is the radiation coefficient for the envelope; and σ is the Stefan-Boltzmann constant. The first term in Eq. (10-30) determines the power delivered as a result of convection, and the second the radiated power.

The geometric factor A may be found from the expression

$$A = 1{,}08 \cdot 10^{-3} \beta \cdot d_б^{\frac{3}{4}} l_б + 0{,}197 \cdot 10^{-3} d_б^2, \qquad (10\text{-}31)$$

where l_b is the length of the envelope, cm; d_b is its diameter, cm; β is a function of the length of the envelope cylinder, as shown in the graph of Fig. 10-27. The radiation coefficient ε is assumed to equal 0.92 for glass and 0.6 for metal envelopes.

A direct solution to Eq. (10-30) is difficult. Thus, in finding the mean temperature of envelopes, it is necessary to utilize the method of successive approximation. The first approximation is obtained

---

* [$P_б$ - P_b - P_{ballon} - $P_{envelope}$.]

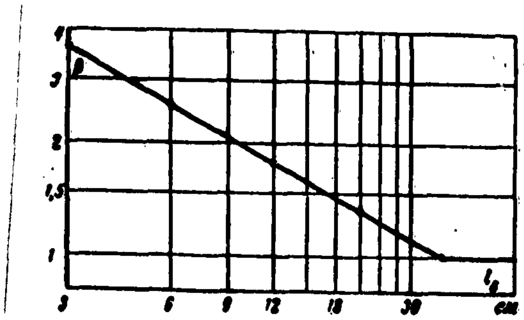

Fig. Fig. 10-27. Graph of the function β.

from the equation

$$T''_б = T'_б - \frac{A(T'_б - T_0)^{\frac{5}{4}} + F_б\varepsilon\sigma[(T'_б)^4 - T_0^4] - P_б}{\frac{5}{4}A(T'_б - T_0)^{\frac{1}{4}} + 4F_б\varepsilon\sigma(T'_б)^3}, \qquad (10\text{-}32)$$

where T'_b is an arbitrarily chosen value of envelope temperature. In practice, it is convenient to choose this value within the 350–400°K range. If the calculation is carried out on the assumption that the ambient temperature is room temperature ($T_0 \approx 300$°K), then for the first approximation, formula (10-32) may be written in simplified form:

$$T''_б = 400 - \frac{316A + 0{,}1F_б\varepsilon - P_б}{0{,}396A + 1{,}49F_б\varepsilon}. \qquad (10\text{-}32a)$$

The second approximation is obtained by substituting the value T''_b into Eq. (10-32) in place of T'_b. In the majority of cases, even the first approximation will give a temperature value that is sufficiently close to the true temperature.

The temperature of glass envelopes should not exceed 120–180°C (390–450°K). The temperature of metal envelopes is limited by the temperature at which the lacquer coating on the envelopes will burn off.

10-6. TUBES FOR ELECTRICAL MEASUREMENTS.

In modern scientific research and in the technology of monitoring several physical processes, it frequently proves necessary to measure direct electric currents of very small values, on the order of 10^{-12}-10^{-15} amp. It is very difficult to make instruments that will yield direct readings of such currents, and thus, such measurements are normally performed by indirect methods. One of the most common and

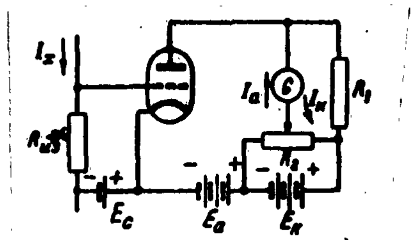

Fig. 10-28. Very simple circuit for measuring small current.

comparatively simple methods is the determination of current value in terms of the voltage drop created by the current across a known very high resistance. The sensitivity of this method of measurement will, evidently, be higher the higher the resistance in the circuit with the measured current.

There are no standard instruments for measuring the difference in potential between the ends of a resistance connected into such a circuit, since the internal resistance of the meter turns out to be considerably lower. Consequently, in order to measure the potential difference created by the measured current, it is necessary to have meters with very high input resistances; these instruments should operate without drawing energy.

A circuit for such an instrument using an electron tube is shown in Fig. 10-28. The ends of the measuring resistor R_{iz}* are connected

* [$R_{\text{из}}$ - R_{iz} - $R_{\text{izmeritel'nyy}}$ - $R_{\text{measuring}}$.]

to the tube input, i.e., to the tube grid, and thus in addition to the constant bias voltage E_c, the voltage across resistor R_{1x} appears on the grid. The magnitude of this voltage will be

$$u_c = I_x R_m, \tag{10-33}$$

where I_x is the direct current being measured.

In the plate circuit of the tube, supplied from the source E_a, the galvanometer G is connected in series with the variable resistor R_2. If resistor R_2 and the internal resistance of the galvanometer are small in comparison with the internal resistance of the tube, as is normally the case, and the plate current is small, then the voltage drop across the series circuit formed by the galvanometer and resistor R_2 may be neglected, and it may be assumed that the tube is operating under static conditions. In the absence of current through the measuring resistance, the value of the plate current will be determined by the plate voltage E_a and the bias E_c. In parallel with resistor R_2 is connected the compensating voltage source E_k, for which R_2 serves as a potentiometer where the galvanometer is connected to this potentiometer through the resistor R_1. The voltage taken from this galvanometer sets up a compensating current in the galvanometer-resistor R_1 circuit, which in the galvanometer has a direction opposite to that of the plate current. By moving the wiper of potentiometer R_2, it is possible to set the compensating current so that it equals the plate current in the absence of current through the measuring resistance. Until current appears through the measuring resistance, the galvanometer needle will indicate zero. Resistor R_1 should be considerably larger than the internal resistance of the galvanometer and resistance R_2 in order for branching of the plate current through the series circuit R_1-R_2 to be negligible.

When current flows through the measuring resistance, the voltage on the grid changes by an amount u_c, and as a result, the plate current changes by an amount ΔI_a, which will be registered by the galvanometer. The magnitude of the plate-current change equals:

$$\Delta I_a = S u_c, \tag{10-34}$$

where S is the transconductance of the tube.

Substituting for u_c from Eq. (10-33), we obtain

$$\Delta I_a = S I_x R_{1z}; \tag{10-35}$$

where the magnitude of the measured current I_x equals:

$$I_x = \frac{\Delta I_a}{S R_{1z}}.$$

Thus, knowing the tube transconductance at the operating point, it is possible to determine the magnitude of the current flowing through the measuring resistance. For example, if the value of the measured current is 10^{-12} amp, the value of the measuring resistance 10^{11} ohms, and the operating-point transconductance 100 μa/v, the change in plate current will equal:

$$\Delta I_a = 100 \cdot 10^{-6} \cdot 10^{-12} \cdot 10^{11} = 10^{-5} a = 10 \text{ мка} \quad *$$

and, consequently, it may be measured with a standard microammeter.

The higher the tube transconductance, the greater sensitivity this circuit will have in measuring small currents. An increase in sensitivity may also be achieved by increasing the value of R_{1z}. The maximum value of the measuring resistance, however, and, consequently, of circuit sensitivity is limited by the tube grid-current value.

---

* [mka - mikroamper - microampere.]

With the circuit just discussed, used to measure small currents, it is assumed that with a constant negative grid potential, the magnitude of the grid current will be zero, or in any case substantially less than that of the measured current. For the majority of common receiving-amplifying tubes, the grid current turns out to be considerably higher than the permissible value. Consequently, for use in circuits measuring small currents, special tubes are required with grid currents not exceeding 10^{-13}-10^{-15} amp. Such tubes are called electric-metering tubes.

In designing electric-metering tubes, it is necessary to know the facters causing grid current to appear, and the numerical values of its components. In general, grid current has many various components, of which the most significant are leakage currents along insulation, electron grid current owing to the initial velocities of electrons, emitted by the cathode, and the electron-emission current from the grid surface. In addition, under specific conditions, there may be a considerable ion current appearing as a result of ionization of residual gases in the tube, and a current of positive ions emitted from the cathode. Research has shown that the two last factors, with a normal vacuum in the tube (10^{-5}-10^{-6} mm Hg) and where the oxide-coated cathode is not at a very high temperature, have no substantial effect upon the magnitude of grid currents.

A decrease in the leakage currents appearing in the presence of a potential difference between the grid and other electrodes, owing to inadequate installation resistance, may be achieved by means of a set of design and production measures. Among these measures are the following:

1. Increasing the resistance of mica insulators by special processing, for example, by means of an alundum coating on their surface.

2. Increasing the leakage paths by designing the insulators with

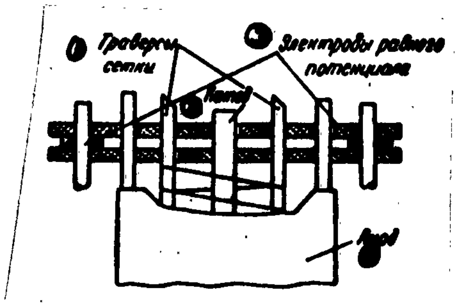

Fig. 10-29. Example of mounting of electrodes of an electric-measuring tube into insulators.
1) Grid traverses; 2) cathode;
3) equipotential electrodes;
4) plate.

notches between mounting holes for electrodes, and by decreasing the perimeter along which the grid is in contact with insulators. The grid lead, as a rule, is brought out through the top of the envelope.

3. Utilization in tube design of the principle of equal potentials, which consists in locating conductors along leakage-current paths; these conductors have a constant potential equal to the grid potential. Fig. 10-29 shows an example of an actual design for mounting the grid of an electric-metering tube, utilizing this principle. The grid is mounted on a separate (external) insulator mounted on a second insulator, in which are placed the cathode and the plate traverses, with the aid of metal traverses and discs to which a potential, equal to the potential of the grid is applied (the lower block of insulators is similar to the upper block).

4. As insulator resistance is impaired basically owing to the deposition of conducting films upon the insulators during production processing of the tubes, it is necessary to reduce the temperature involved in heat treatment of tube parts, and primarily as regards the cathode. Thus, cathodes are utilized in electric-measuring tubes at lowered temperatures.

The measures discussed makes it possible to decrease the leakage currents to the values required, on the order of 10^{-14}-10^{-15} amp. In many cases, it turns out that it is enough to carry out only some of the measures described.

The magnitude of the electron grid current depends upon cathode temperature, which determines the distribution of the initial velocities of thermal electrons, and the grid potential, allowing for contact potentials. At a cathode temperature that is not too high (800-900°K) with adequate retardation of the electrons emitted by the cathode, there is an adequate bias on the grid, on the order of -1.5 to -2.0 v.

Very great problems are caused by electron emission from the surface of grids. The basic type of grid electronic emission is photo-emission, caused, first, by radiation from the heated cathode and, second, by extremely soft X-rays appearing upon retardation of the electrons arriving at the plate. Thermionic emission from the cold grids of electric-metering tubes is practically nonexistent.

The magnitude of photoelectric emission depends upon the intensity of the radiation, its spectral composition, and the surface area and surface properties of the grid. The intensity and spectral composition of the radiation of the cathode is determined by the cathode temperature. A decrease in temperature to a value on the order of 800-900°K, with a sufficiently small grid surface will decrease the photocurrent resulting from cathode radiation to a value on the order of 10^{-14}-10^{-15} amp. It is not possible to drop the temperature further, since in this case there will be a sharp drop in cathode emission.

The intensity of very soft X-rays from the plate will depend upon the magnitude of the plate current and plate voltage of the tube. Theoretical and experimental investigations carried out by Yu. F. Zarutskiy have led to the following equation:

$$I_\Phi = cI_a Z(U_a - \varphi_c)^4, * \qquad (10\text{-}36)$$

where I_f is the photocurrent; I_a is the plate current of the tube, which causes the very soft X-rays; U_a is the plate voltage; Z is the atomic number of the plate material; φ_c is the grid work function; c is a geometric factor that depends upon the dimensions of the grid-plate system.

Since the values of plate current, plate voltage, and the minimum necessary negative bias are directly connected with the tube characteristics – the transconductance and the amplification factor, there is also a direct connection between the characteristics of an electric-metering tube and the value of grid current. In other words, the minimum permissible value of grid current, determined basically by grid photoemission, associated with the very soft radiation from the plate, depends upon the tube characteristics.

It is clear from Eq. (10-36) that in order to decrease the grid photoemission, it is primarily necessary to drop the plate voltage. Thus, electric-metering tubes operate, as a rule, at plate voltages below 10 v. In order to obtain an adequate transconductance at such low plate voltages, it is necessary to design tubes with low amplification factors. Thus, electric-metering triodes have penetrances close to one. An increased transconductance with high gain may be obtained in a tetrode.

In electric-metering tubes, in order to obtain the required value of transconductance, the tetrode-with-cathode-grid construction is employed. The operating principle of such tubes consists in the fact that there is a single grid between the control grid and the cathode, at a positive voltage equal to or somewhat less than the plate potential.

* [I_Φ - I_f - $I_{fototok}$ - $I_{photocurrent}$.]

TABLE 10-3.

Обозначение лампы	Тип лампы	U_a, в	E_c, в	$I_{c.\,макс}$, а	S, мка/в	μ
ЭМ-3	Тетрод с катодной сеткой	6	—3	10^{-13}	25	1,4
ЭМ-4	Триод	8	—1,7	$8\ 10^{-16}$	80	2,2

1) Tube designation; 2) type of tube; 3) U_a, v; 4) E_c, v; 5) $I_{c.\,max}$, amp; 6) S, μa/v; 7) EM-3; 8) EM-4; 9) tetrode with cathode grid; 10) triode.

The presence of the positive grid leads to a decrease in the absolute value of the potential minimum in the space-charge region near the cathode, which in turn provides an increase in plate currents and tube transconductance. The effect of the control grid in an electric-metering tetrode with a cathode grid is similar to the action of a pentode suppressor grid in a dual-control pentode. A change in control-grid potential causes a change in plate current owing to redistribution of cathode current between the plate and the cathode grid.

The utilization of control grids having small grid-wire diameters with large penetrances makes it possible to bring the control grid near to the cathode, and as a result, to obtain an increase in diode transconductance. On the other hand, the presence of a positive cathode grid in tetrodes with cathode grids increases tube grid current owing to additional very soft radiation from it. Thus, designing of electric-metering tetrodes with cathode grids is evidently not very promising.

Thus, the basic structural features of electric-metering tubes are low cathode temperature, which is possible only where oxide-coated cathodes are utilized; insulator construction providing low leakage currents, low plate voltages, and small control-grid surface.

The characteristics of certain types of domestic electric-metering

tubes are given in Table 10-3.

10-7. TUBES UTILIZING SECONDARY EMISSION OF ELECTRONS.

One of the most promising methods for increasing the transconductance of receiving-amplifying tubes is the utilization of secondary electron emission. Let us consider the processes occurring in tubes with secondary-electron emission, for example, the tube whose arrangement is shown in Fig. 10-30. The tube has an indirectly heated cathode and two grids: a control grid and a screen grid. The stream of electrons, leaving the second grid, are directed toward an electrode called a dynode, to which is applied a high positive potential (on the order of 150 v). When a stream of electrons moves to the dynode, it is focused with the aid of focusing traverses, connected with the cathode. This focusing of the stream is necessary in order for the primary beam of electrons not to fall directly on the anode, which is shaped as a flat plate and located within a dynode.

The inside surface of a dynode is treated in order to increase its secondary-emission coefficient; this process should provide adequate stability of the dynode in time under extended electron bombardment, and under the action of contaminants leaving other electrodes in the form of dust. Materials for manufacturing dynodes may be various secondary-electron emitters having secondary-emission coefficients on the order of 2.5-3.5.

The stream of electrons, arriving at a dynode, causes emission of secondary electrons. Since the anode potential is about 100 v higher than the dynode potential, the secondary electrons move toward the anode, forming an anode current. The dynode, thus, becomes a source of electrons, creating a space charge between the dynode and the plate. Where the anode voltage is close to the dynode potential,

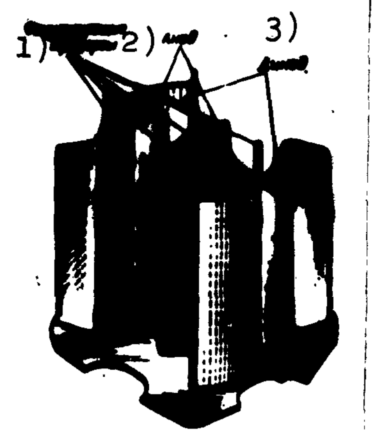

Fig. 10-30. A tube utilizing secondary electron emission. 1) Focusing traverses; 2) plate; 3) dynode.

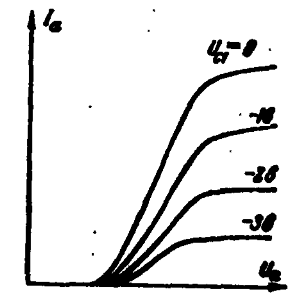
Fig. 10-31. Family of plate characteristic curves for a tube with secondary electron emission.

this space charge sets up a potential minimum in the plate-dynode space, which leads to partial return of secondary electrons to the dynode. An increase in anode voltage causes a rapid rise in anode current owing to resorbtion of the space charge. When the anode current becomes equal to the secondary-emission current from the dynode, there is no further rise in it as anode voltage goes up. Figure 10-31 shows an example of a family of anode characteristics for a tube with secondary-electron emission. Where the anode voltage is lower than the dynode potential, there is no anode current, since the secondary electrons are returned to the dynode by the electric field set up in the dynode-anode space.

If we designate the current flowing to the dynode from the cathode by I_1, then the magnitude of the secondary-emission current, which with sufficient voltage on the anode equals the anode current, may be found from the following expression:

$$I_2 = I_a = \sigma I_1. \qquad (10\text{-}37)$$

61-102

The dynode current is defined as the difference between the primary-electron current to the dynode and the anode current:

$$I_д = I_1 - I_a = I_1(1-\sigma).* \qquad (10\text{-}38)$$

Where the secondary-emission coefficient σ is larger than 1, the dynode current is negative (provided that the anode voltage is high enough).

Figure 10-32 shows sample grid characteristic curves indicating the dependence of the anode current and the dynode current upon first-grid voltage with constant voltages on the second-grid, dynode, and anode. As is clear from the curve, as the first-grid voltage increases, there is a simultaneous rise in anode and dynode currents.

The anode-grid transconductance of the tube depends not only upon the geometry of the cathode and first grid and the current division, as is the case for a tetrode, but also upon the secondary-emission coefficient of the dynode surface:

$$S_a = \frac{dI_a}{dU_{c1}} = \sigma\frac{dI_1}{dU_{c2}} = \sigma S_0, \qquad (10\text{-}39)$$

where S_0 is the tetrode transconductance of the tube, indicating the effect of a change in first-grid potential on the magnitude of the primary stream of electrons. Thus, utilizing secondary emission, it is possible to obtain an increase in transconductance in comparison with a tetrode of similar design by a factor of σ.

In like manner, we may find the transconductance for the dynode-grid system:

$$S_д = \frac{dI_д}{dU_{c1}} = (1-\sigma)S_0. \qquad (10\text{-}40)$$

---

* [$I_д$ - I_d - I_{dinod} - I_{dynode}.]

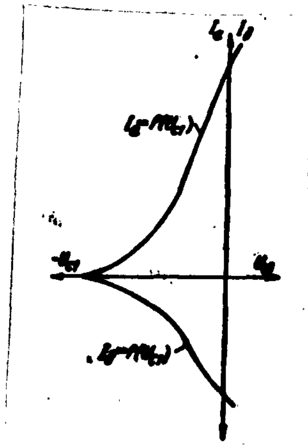

Fig. 10-32. Plate-grid characteristic curves of tube secondary electron emission.

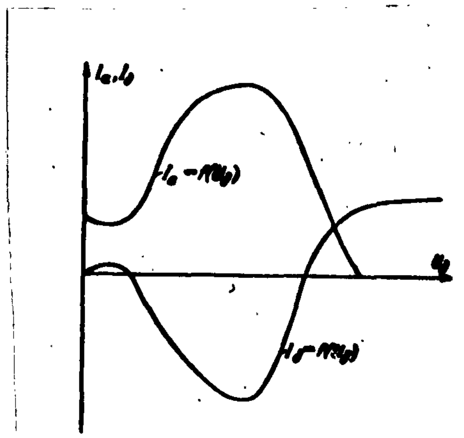

Fig. 10-33. Dynode characteristic curves of tube with secondary electron emission.

In order to obtain the best tube characteristics, it is extremely important to make a proper choice of dynode potential. Figure 10-33 shows the dynode characteristic curves for anode current and dynode current $I_a = f(U_d)$ and $I_d = f(U_d)$, taken with constant potential on the remaining electrodes. Maximum dynode and anode currents are obtained at dynode voltages equal to about 60 per cent of the anode voltage. With a decrease in dynode voltage, the absolute values of anode and dynode currents drop, first, owing to a change in the current distribution between the second grid and the dynode-anode system, and second, owing to defocusing of the primary-electron stream, as a result of which there is a large number of primary electrons that fall directly upon the anode and, finally, owing to a drop in the secondary-emission coefficient attendant upon the decrease in primary-electron energy. For dynode voltages lower than 40-50 v, the secondary-emission coefficient is less than 1, and the dynode current proves to be positive. In this case, the magnitude of the anode current is determined solely by the number of primary electrons arriving at the anode. As the dynode

voltage rises above the optimum value, the absolute values of anode and dynode currents drop as a result of the partial return of secondary electrons to the dynode, and when the dynode voltage becomes equal to the anode voltage, movement of secondary electrons to the anode ceases, the anode current becomes equal to zero, while the dynode current becomes positive, and equal to the primary-electron current.

In tubes utilizing secondary-electron emission, it has been possible to obtain a high transconductance on the order of 25-30 ma/v, with a relatively small cathode current. In such tubes, it has been possible to obtain a ratio $S/C_{in} + C_{out}$ on the order of 2-2.5 ma/v·$\mu\mu$f.

A substantial advantage of tubes utilizing secondary emission is a relatively small amount of power dissipated on the anode with high anode current. The reason for this is that the secondary electrons creating the anode current arrive at the anode with an energy produced not by the anode voltage, but by the difference of potential between the anode and dynode. Consequently, the power dissipated on the anode in such a tube equals:

$$P_a = I_a (U_a - U_d). \qquad (10\text{-}41)$$

The basic drawbacks to tubes utilizing secondary-electron emission are relatively short life, caused by a decrease in secondary-emission coefficient of the dynode surface under the action of extended electron bombardment, and high noise level.

10-8. ELECTRON-BEAM TUNING INDICATORS.

Precision of tuning for radio-receiver tuned circuits may be monitored with the aid of so-called electron-beam tuning indicators. The common electron-beam indicator takes the form of a combination of a triode with an electron-beam system, as illustrated in Fig. 10-34. A phosphor is applied to the inner surface of a truncated cone; this

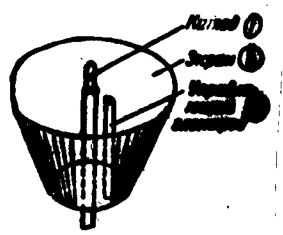

Fig. 10-34. Electron-beam system of tuning indicator. 1) Cathode; 2) screen; 3) control electrode.

is a material that glows when electrons strike it. Willemite (Zn_2SiO_4) is normally utilized as the phosphor in indicator tubes; it gives off a green glow. A cathode is located on the axis of the cone, while between the cathode and the cone a control electrode, in the form of a rod, is installed. The cone is maintained at a high positive potential. Under the action of the electric field, electrons from the cathode bombard the exposed phosphor of the inner cone surface, causing it to glow.

If the control electrode remains neutral or is kept at a potential equal to the potential of the neighborhood in which it is located, the presence of the control electrode will not distort the electric field in the space between the cathode and the cone. As a result, the electric-current density proves to be the same over the entire surface of the cone, and the illumination will be uniform. Only in the region of the surface lying under the control grid will the illumination be weakened, since a portion of the electrons will be intercepted by the control electrode (Fig. 10-35a). If the potential of the control electrode is changed, the symmetry of the field is destroyed, and the trajectory of the electrons moving near the control grid will be bent. Where the control-grid potential is decreased, the electron trajectories will

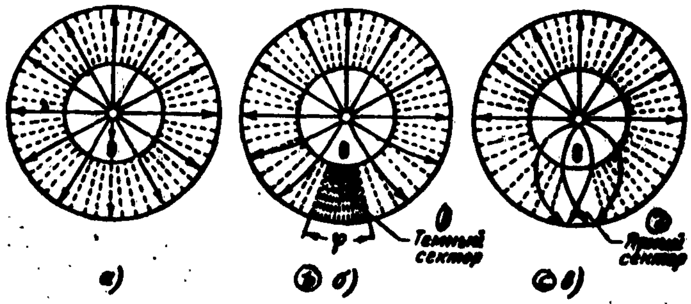

Fig. 10-35. Change in electron trajectories with a change in control-electrode potential. 1) Dark sector; 2) bright sector.

move away from it and electrons will in general not strike that portion of the cone surface. On this portion of the screen, a dark sector will appear (Fig. 10-35b). Conversely, when the control-grid potential is raised, the trajectories of the electrons are bent toward it, and more electrons will strike the section lying underneath the grid than will strike other sections of the screen. In this case, a sector will appear on the screen that is more brightly illuminated than the remaining portion of the screen (Fig. 10-35c).

In an electron-beam indicator, the control electrode is connected to the triode plate, and its potential, consequently, will be called the plate potential. The electron-beam system and the triode have a common cathode. The arrangement of the 6E5C tuning indicator is shown in Fig. 10-36. The electron-beam system is in the upper part of the tube, and it is observed through the transparent top of the tube envelope.

Figure 10-37 shows a circuit used with the indicator. The alternating signal voltage is rectified by a diode circuit and applied from the load resistor to the grid of the triode of the electron-beam indicator. A positive voltage from a source, E_a, is applied to the screen

(cone); the triode plate is connected with the power supply through a load resistor having a value of 1-2 megohm. When there is current in the plate circuit of the triode, the magnitude of the plate voltage will be less than the screen voltage by the amount of the voltage drop across the load resistor. Evidently, the larger the triode plate current, the larger the difference between the potentials of the screen and plate. Since the control electrode is connected to the

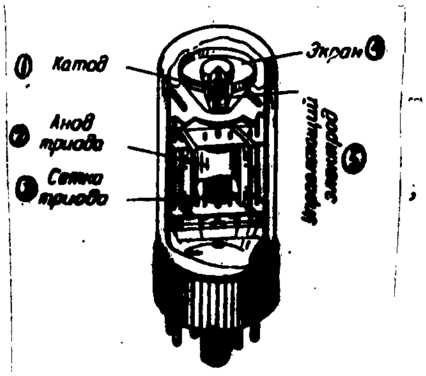

Fig. 10-36. 6E5C electron-beam indicator.
1) Cathode; 2) triode plate; 3) triode grid; 4) screen; 5) control electrode.

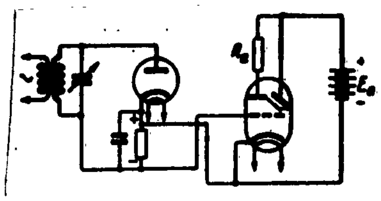

Fig. 10-37. Circuit for utilizing an electron-beam indicator tube.

triode plate, a decrease in plate current of the triode leads to an increase in control-electrode potential and, consequently, to a decrease in the dimensions of the dark sector on the screen.

When a tuned circuit is in resonance with the frequency of a signal being received, the RF voltage in the tuned circuit is at a maximum, and the rectified voltage will also be at a maximum. Since a negative voltage is applied to the triode grid, accurate tuning of the circuit will result in minimum grid potential. In this case, the potential of the triode plate and the control electrode will be at its maximum, which corresponds to minimum size of the dark sector. Evidently, when tuning through a weak signal, there will only be a decrease in the size of the dark sector at the instant of precise

tuning, while when tuning through a strong signal, a brighter sector

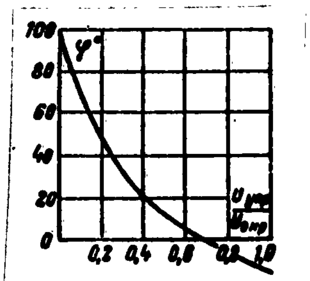

Fig. 10-38. Graph of $\varphi = f(U_{upr}/U_{ekr})$* for 6E5C tuning indicator.

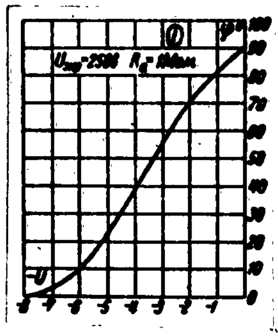

Fig. 10-39. Characteristic curve $\varphi = F(U_c)$ for 6E5C tuning indicator.
1) Megohm.

will appear.

The magnitude of the vertex angle of dark sector on the indicator screen depends upon the relationship between the voltages on the screen and control electrode, or, with a fixed screen voltage, upon the control-electrode potential. Figure 10-38 shows a graph of the dark-sector angle as a function of the magnitude of the ratio of control-electrode voltage to screen voltage. Under dynamic operating conditions for the indicator triode, the plate voltage, equal to the control-electrode voltage, will be a function of the triode-grid voltage and, consequently, the dark-sector angle will also be a function of the grid voltage. The characteristic curve $\varphi = f(U_c)$ for the type 6E5C indicator is shown in Fig. 10-39.

* [$U_{\text{упр}} = U_{upr} = U_{upravlyayushchiy} = U_{control}$.]
[$U_{\text{экр}} = U_{ekr} = U_{ekran} = U_{screen}$.]

Chapter Eleven

TRANSMITTING TUBES

11-1. CLASSIFICATION AND DESIGNATIONS FOR TRANSMITTING TUBES.

The operating peculiarities of transmitting tubes and the requirements applicable to them are responsible for the substantial structural differences between transmitting tubes and receiving-amplifier tubes. The most important requirement of transmitting tubes is that they deliver a large amount of RF power at as high a circuit efficiency as possible. The RF power delivered by the tube, the tube function, and the maximum working frequency may serve as criteria for tube classification.

With a given plate-circuit efficiency, the RF power that may be obtained from a transmitter circuit is proportional to the power dissipated by the plate of the tube:

$$P_{\kappa} = P_{a} \frac{\eta}{1-\eta}.$$

Consequently, the possibility of obtaining the required RF power from a given tube may be characterized by the power dissipated on the plate of the tube.

On the basis of power, all transmitting tubes may be classified into the following groups:

a) Low-power tubes, in which the permissible continuous plate dissipation does not exceed about 50 watts. Plates of these tubes are naturally cooled, i.e., heat is carried off to the surrounding space as a result of heat radiation. The gross structural features of low-power transmitting tubes are in the majority of cases similar to these of receiving-amplifying tubes (Fig. 11-1a, b, c).

b) Medium-power tubes, with a continuous plate dissipation

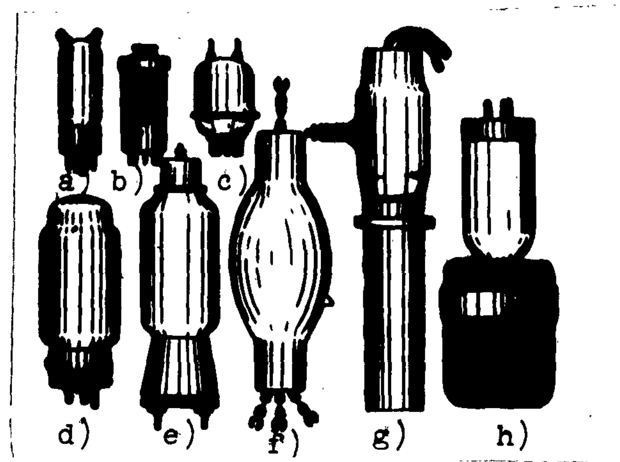

Fig. 11-1. External appearance of transmitting tubes.
a), b), and c) Low-power transmitting tubes; d), e) and f) medium-power transmitting tubes; g) high-power water-cooled transmitting tube;
h) high-powered air-cooled transmitting tube.

of up to 1000 watts (Fig. 11-1d, e, f). In these tubes, as in the low-power tubes, the plates are located within a glass envelope, and are cooled naturally. The increase in plate power dissipation, as well as the utilization of high voltages, leads to the necessity of increasing tube dimensions, and of making major changes in lead, insulator, and envelope design in comparison with the low-power transmitting tubes.

c) High-power tubes, in which the permissible continuous plate dissipation exceeds 1000 watts (200 kw and above). Exterior plates of such tubes are cooled either with air, blown past radiating fins attached to the plate (Fig. 11-1h), or with water that flows over a surface of the plate (Fig. 11-1g).

In accordance with the maximum working frequency, transmitting tubes may be classified into low-frequency tubes, working in audio-amplification circuits, RF tubes, designed to work in oscillators and

61-107

amplifiers handling radiofrequencies (up to 30 Mc) and SHF tubes, whose features are discussed in the following chapter.

The working-frequency region is a distinguishing feature of transmitting-tube designations. A tube designation consists of three elements. The first element is represented by letters that specify the field of application of the tube: GK-transmitting tubes for the long- and short-wave bands (upper frequency limit of up to 25 Mc); GU-transmitting tubes for VHF (upper frequency limit of up to 600 Mc); GS-transmitter tubes for the centimeter range; GM-tubes operating in low-frequency (audio) power-amplifier circuits; GMI-similar to previous type, but designed for pulse operation; GI-high-frequency pulse-generating tubes.* The numbers forming the second element in a designation are separated from the first element by a hyphen, and indicate the sequence number of the given type of device. Transmitting tubes with forced air or water cooling have a third element in their designation, indicating the type of cooling: the letter A is used for water-cooled tubes, and the letter B (Б) for air-cooled. Thus, the designation GK-1A corresponds to a water-cooled RF tube, the GU-10B is an air-cooled VHF tube, the GM-60 is an audio tube with natural cooling of the plate (the absence of a third element in the designation indicates that the plate is cooled solely by radiation).

---

* [GK - generator, korotkiy - transmitting, short (wave)]
[GU - generator, ul'trakorotkiy - transmitting, ultra-shortwave/VHF]
[GS - generator, santimetrovyy - transmitting, centimetric]
[GM - generator, moshchnostnyy - transmitting, power]
[GMI - generator, moshchnostnyy, impul'snyy - transmitting, power, pulse]
[GI - generator, impul'snyy - transmitting, pulse].

11-2. FURTHER PARAMETERS OF TRANSMITTER TUBES.

As in the case of receiving-amplifying tubes, the possibility of utilizing transmitting tubes in various circuits is determined not only by the static characteristics, but by many additional properties as well. Considerations that have been discussed in connection with the additional characteristics of receiving-amplifying tubes also apply to low-power transmitting tubes. The peculiarities and specifications applicable to medium-power transmitting tubes and to high-power tubes are connected basically with the high interelectrode voltages (both AC and DC), and with the high values of electrode currents.

a) Magnitude of emission current. In contrast to receiving-amplifying tubes, the majority of transmitting tubes for medium and high powers operate with an excursion in the saturation region. Consequently, the peak plate current is limited by the cathode emission. It is undesirable to design tube cathodes with an emission reserve, since in this case, there is considerable increase in the heating power, which is quite substantial in transmitting tubes. Thus, the cathode emission current is normally taken as one of the characteristics of transmitting tubes.

The cathode emission has special importance in pulse-generator tubes, in whi the plate current flows during a short pulse. In this case, the power in a pulse is completely determined by the size of the pulsed cathode emission, which is a characteristic of pulse-generating tubes.

In tubes with tungsten and carburized cathodes, the emission current depends to a great degree upon production treatment of the cathodes in assembling the tubes, which should be subjected to careful quality control.

b) Power dissipated on grids. The majority of transmitting tubes

operate with an excursion in the region of high positive grid potentials and, consequently, with large grid currents. In this case, considerable power is dissipated on the grids. A great deal of power is also dissipated on the screen grids of tetrode and pentode transmitting tubes. In many cases, maximum RF power that is obtained in a vacuum tube circuit is limited not by the maximum permissible power dissipated on the plate, but by the power dissipated on the grids.

It should be noted that the magnitude of grid current measured by a meter connected into the grid circuit proves to be less than the primary-electron current of the grid as the result of the large secondary emission of electrons from the grid. The power liberated at a grid depends upon the primary-electron current of the grid and its voltage (the relative proportion of the power that is associated with the secondary electrons is negligible). Thus, in calculating the power liberated at a grid, it is necessary to take into account the actual value of the primary-electron current for the grid.

In many transmitting tubes, the amount of power dissipated on grids is an important operating characteristic.

c) Grid back currents. As a result of the heating up of transmitting-tube grids, there may appear thermionic emission from grids and this may reach considerable magnitudes. Since during the negative half cycle of the alternating voltage on the grid, the difference of potential between grid and cathode may reach values on the order of several hundreds of volts, or even a kilovolt; intense electronic bombardment of the cathode may occur, leading to damage to the cathode surface. The appearance of thermionic grid emission in tubes with oxide-coated and carburized cathodes is especially dangerous.

Thus, a decrease in grid thermionic emission is one of the most important problems involved in transmitting-tube design. In low- and

medium-power transmitting tubes, the grids are coated with gold, platinum, copper, or blackened in order to increase the surface radiation coefficient. As has been noted, blackening of a grid is desirable only in tubes having open-structure-type plates, or in tubes with relatively low plate temperatures. Otherwise, blackening may lead to the opposite result, since in this case there is an increase in the absorption of power radiated by other, hotter electrodes.

In high-power tubes, the grids are normally coated with zirconium, which not only provides absorptive grid surfaces, but also raises the work function and the coefficient of radiation. Sometimes, inner surfaces of power-tube plates are blackened (by black chrome plating), which decreases reflection of radiation by the plate and, consequently, improves grid-cooling conditions.

d) Cold-cathode resistance. In transmitting tubes with tungsten and carburized cathodes, the cold-cathode resistance differs sharply from the working-temperature resistance. Thus, the initial cathode current may be very large. For example, in the GK-3A triode oscillator, at rated heating current of 430 amp, the cold-cathode resistance is less by a factor of 11 than the working-temperature resistance. This means that if the working heating voltage is applied in the initial instant, the starting current will equal nearly 5000 amp.

Large starting currents are also impermissible in view of the fact that destructive electrodynamic forces appear between branches of the cathode. Thus, the resistance of the unheated cathode and the maximum initial heating current are used as characteristics for transmitting tubes. This makes it possible to select the correct value of heating voltage for application to the tube when it is first turned on. The initial heating current should not exceed, for the majority of tubes, 150 per cent of the rated heating current.

e) Maximum permissible interelectrode voltages. Arc-overs may

appear between electrodes during tube operation, where high potential differences are involved. The reasons for the appearance of arc-overs may be deterioration of the vacuum in the tube, secondary emission from parts (including mounting elements) and deterioration of the insulating properties of tube internal insulation.

When interelectrode arcing occurs, it is possible that tubes will be damaged, especially if arcing appears at points where leads or plates of the tube are sealed into glass. Thus, the values of interelectrode voltages, even instantaneous values, should not exceed the values established for each tube.

f) Mechanical and climatic tube characteristics. As in the case of receiving-amplifying tubes, transmitting tubes may be subjected to various mechanical effects. Mechanical effects are especially dangerous to medium- and high-power transmitting tubes owing to the large weight of the elements and large linear dimensions, which result in considerable forces at points where parts are fastened when external forces are applied – vibration, shocks, etc. Thus, destruction of joints between tube elements and deformation of elements is more probable in transmitting tubes than in receiving-amplifying tubes. In addition, the long linear dimensions of elements causes a decrease in internal resonant frequency to hundreds of cycles. Thus, many high-power transmitting tubes are designed to operate only in stationary installations, with special measures being taken to guard against external mechanical effects.

In the process of transporting the tubes themselves, and the devices in which they are utilized, however, mechanical influences may make themselves felt. Thus, the vibration resistance is sometimes given for many transmitting tubes, i.e., the ability to retain the initial parameters after vibration.

For high-power transmitting tubes, and many types of medium-power tubes, changes in climatic conditions under which tubes operate may prove dangerous. The high powers dissipated by tubes and envelopes frequently necessitate special cooling conditions. Overheating of points where metal is sealed into glass, which may lead to tube damage, is especially dangerous. The permissible temperatures for external plates, seals, and envelopes are normally specified as operating characteristics. Elevated humidity may lead to arcing-over along external surfaces of envelopes and mounts, and a decrease in atmospheric pressure to the appearance of a high-frequency discharge.

11-3. STRUCTURAL FEATURES OF TRANSMITTING TUBES.

The structural features of transmitting tubes are determined by tube function, the magnitude of the power dissipated on the plate, the maximum working voltages between electrodes, and the working-frequency range. In addition, tube construction is related to their operating conditions.

General requirements applicable to all types of transmitting tubes are high cathode-emission power, adequate for the given operating conditions, high plate and grid dissipation, and high electrical strength of interelectrode insulation, which permits tubes to be operated at high interelectrode voltages.

Transmitting tubes designed for operation in oscillator circuits and RF power-amplifier circuits have characteristic curves shifted to the right, which, as has already been mentioned, makes it possible to operate these tubes with high plate efficiency under class C conditions, with a small amount of grid bias. For triodes, a right-hand shifted characteristic curve corresponds to high amplification factor, while for triodes and pentodes, it corresponds to low control-grid penetrance

and relatively low DC voltage on the screen grid.

In order to increase the plate-voltage utilization factor, it is desirable that the critical operating line have large slope, and from the point of view of decreasing the power dissipated on the grid, it is necessary to increase the current-division factor.

Low-power transmitting tubes, designed to operate in RF oscillaters (up to 20-25 Mc), are normally pentodes or beam-power tetrodes. The advantage of these types of tubes lies in the possibility of obtaining high power gain per stage of amplification, and good frequency response in comparison with triodes. The same is true of tubes for medium-power applications (maximum plate dissipation up to 500 watts). The limitation of the possibility of utilizing pentodes where tube power is further increased is associated with difficulties appearing owing to an increase in screen dissipation. The use of a beam system for the control and screen grids (beam pentodes) yields a certain decrease in screen dissipation, and a drop in screen-grid potential. Screen voltage in pentode and beam-tetrode transmitters normally amounts to 20 to 60 per cent of the plate voltage.

Transmitter tubes dissipating more than 1 kw on the plate are, in the majority of cases, triodes. The only exceptions are certain VHF tetrodes dissipating up to 5-10 kw on the plate. In such tetrodes, the screen-grid voltage amounts to 25-25 per cent of the plate voltage.

Transmitter tubes for low-frequency power amplification (modulator tubes) are, as a rule, of the triode type. The gain of such tubes, depending upon the plate voltage and rate of power, may vary from 5 to 20, which corresponds to a "left-hand" position of the plate-grid characteristic curves. Such triodes are operated class A; in the majority of cases there is an excursion in the positive grid-potential region and, consequently, there is an excursion for the grid current. Triodes

designed to operate without grid currents have low gains — on the order of 5-7. When there is an excursion in the positive grid-potential region, the grid currents are relatively small and, accordingly, the power dissipated on the grids is also small in comparison with RF transmitters.

Tubes used for power amplification of brief pulses, operating in pulse-modulator circuits, form a separate group. As an example of the application of such tubes, we may consider the operation of a tube used in a pulse-modulator circuit with partial capacitor discharge (Fig. 11-2).

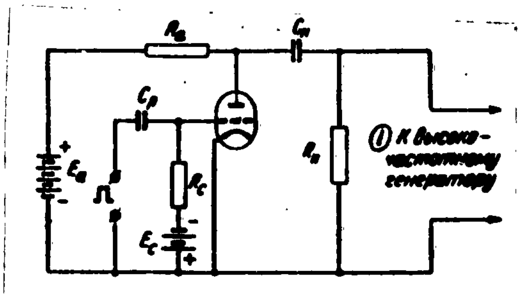

Fig. 11-2. Pulse-modulator circuit with partial capacitor discharge.
1) To RF oscillator.

In this circuit, a bias voltage E_c, cutting the tube off, is applied to the grid. Thus, in the absence of an external voltage pulse, the current through the tube is zero and, consequently, the plate voltage equals the plate-supply voltage E_a, since the voltage drop across the resistor R_a connected into the plate circuit is zero. In this case, the integrating capacitor C_n* is charged to a voltage equal to the plate voltage. The amount of energy stored by this capacitor is proportional to its capacitance and to the voltage across it:

* [C_H - C_n - $C_{nakopitel'nyy}$ - $C_{integrating}$.]

$$W_c = \frac{C_n U_c^2}{2}. \qquad (11\text{-}1)$$

The modulator output is applied to the RF transmitting tube, and the modulator serves as the source of plate voltage for the transmitting tube. When the tube is cut off, the voltage across the modulator output is zero, and the oscillator will not operate.

If a positive voltage pulse is applied to the grid of the modulator tube, the tube is rendered conducting, and current flows through it. Capacitor C_n begins to discharge through the tube and through resistor R_n, and through the transmitting tube connected in parallel. A current appears in the plate circuit equal to the difference between the plate current and the capacitor-discharge current, and the plate voltage drops sharply owing to the voltage drop across resistor R_a. Normally, the circuit constants are so chosen that the tube plate voltage when operating is $U_{a.\ min} = 0.1E_a$. Then a voltage will appear in the plate circuit of the RF oscillator for a period of time equal to the pulse duration; the voltage is classified as the difference between the voltage across the capacitor C_n and the plate voltage:

$$U_{a.r} = U_c - U_{a.мин} \approx 0{,}9E_a.^{*} \qquad (11\text{-}2)$$

The value of capacitor C_n is so chosen that after a period of discharge equal to the pulse length, the voltage across the capacitor will fall slightly (by 2-5 per cent), and this voltage may be assumed to be constant for all practical purposes.

When the pulse ceases to act, the tube is again cut off, and capacitor C_n begins to discharge until the voltage across it becomes equal to E_a.

* [$U_{a.r}$ - $U_{a.g}$ - $U_{anod.\ generator}$ - $U_{plate\ oscillator}$.]

Thus, during the time between pulses, the capacitor stores energy from the modulator plate-supply source, releasing it during the pulse.

The power expended in the plate circuit of an oscillator is determined by the magnitude of the RF power and the oscillator plate-circuit efficiency:

$$P_{a} = \frac{P_{к}}{\eta}. \tag{11-3}$$

Then, in the oscillator circuit, the current drawn equals:

$$I_{0г} = \frac{P_{a}}{U_{a.г}}. \tag{11-4}$$

When the capacitor discharges through the tube, the discharge current is distributed between the load (oscillator) and resistor R_n:

$$I_{c.н} = I_{0г} + I_{Rн},^{*} \tag{11-5}$$

and the magnitude of the tube plate current, which equals the sum of the discharge current and the current in the plate circuit, is found from the expression

$$I_{a} = I_{c.н} + I_{Rш} = I_{0г} + R_{Rн} + I_{Rш}. \tag{11-6}$$

The total of the currents through resistors $R_ш$ and R_n normally runs from 7 to 12 per cent of the plate current. Consequently,

$$I_{a} = mI_{0г}, \tag{11-7}$$

where m = 1.05-1.15.

The amount of current dissipated on the tube plate over the length of the pulse equals:

---

* [н - n - nagruzka - load.]

$$P_{a.u} = I_a U_{a.\text{max}} \tag{11-8}$$

from which it is possible to determine the mean power dissipated on the tube plate:

$$P_a = \frac{P_{a.u}}{Q}, \tag{11-9}$$

where Q is the pulse duty factor.

Thus, pulse modulator tubes should operate at very high plate voltages in the absence of a pulse, and deliver large amounts of power during the pulse, which corresponds to a high value of pulse plate current. We should note that when a voltage pulse is applied to the grid on the order of hundreds and thousands of volts, no space charge is set up near the cathode, and the value of current from the tube cathode is determined solely by the cathode emission.

It is clear that the basic characteristics of a pulse modulator tube are the pulse plate current, the maximum permissible voltage between plate and cathode with the tube cut off, determined by the strength of the insulation, and the permissible plate dissipation. In addition to the power dissipated on the plate, a considerable amount of power is dissipated by the grids of the tube, which to a large degree limits the power delivered by the tube.

Medium-power pulse modulator tubes, as a rule, take the form of beam-power tetrodes; in these tubes, the screen voltage amounts all-in-all to 5-8 per cent of the maximum plate voltage, on the order of 20-30 kv. A relatively small screen voltage decreases the absolute value of the plate-current cutoff voltage and, consequently, in these tubes, the bias voltage is dropped as is the required voltage-pulse amplitude for the control grid.

The amount of power dissipated on the plate of the tube is the

primary factor affecting tube dimensions. All of the power radiated by the electrodes of low-power and medium-power tubes passes through the glass of the envelope and, consequently, the surface of the envelope should be great enough to prevent overheating of the glass.

The magnitude of the working voltages between the electrodes determines the arrangement of leads and the construction of the tube stems. At high interelectrode voltages (above 700-800 v), leads for these electrodes are brought through separate stems, which are sealed into necks in the envelope, and are separated from each other.

The maximum working frequency has a substantial influence on transmitting-tube design. Tubes for the long-wave band normally use tubular stems with long leads and mounting brackets for the electrodes. Grid leads for such tubes are sealed into the cathode stem (Fig. 11-1f). With tubes operating in the short-wave range (from 3 to 25 Mc), it is better to use button or disc stems and short leads; in this case, the control-grid leads are separated from the plate and cathode leads.

11-4. TRANSMITTING-TUBE CATHODES.

The choice of cathode type and design for a transmitter tube is determined by the required cathode emission, by the working voltages of the tube, and by the temperature of the elements surrounding the cathode.

The best emission figures are given, as is known, by an oxide-coated cathode; this type of cathode is widely employed in low-power and medium-power transmitter tubes, and in the majority of tubes for pulse work (with the exception of high-power tubes). The application of the oxide-coated cathode is limited, however, by the possibility of its damage owing to bombardment by ions of residual gases in the presence of high interelectrode voltages, and also as a result of

overheating of the cathode owing to radiation from the inner surface of the plate when it is at a high temperature. Another limitation in the employment of oxide-coated cathodes in tubes where the surrounding electrodes are at high temperatures (more than 500-600°C) results from the possibility of the occurrence of thermionic emission from these electrodes as a result of deposition upon them of fine particles of alkali-earth metals and of oxide from the cathode.

Thus, the oxide-coated cathode is utilized, as a rule, in transmitter tubes that are in continuous service with a continuous plate dissipation of roughly up to 100 watts, and in medium-power tubes for pulse operation with up to 150 watts plate power. In pulse tubes, designed for large cathode currents during the pulse, synthetic-oxide-coated cathodes are frequently used; they differ from those commonly used in that on the surface of the cathode base-metal layer, and throughout the oxide-coating layer, a fine nickel powder is introduced, sintered with the base material and the oxide. The presence of the pure metal inclusion in the oxide layer decreases its resistance and increases the contact with the base-layer surface, resulting in a drop in the heating of the oxide layer by the emission current and, consequently, an increase in the permissible current density through the oxide layer.

Oxide-coated cathode construction for transmitting tubes is shown in Fig. 11-3. Unlike receiving-tube cathodes, transmitting-tube cathodes are insulated from their heaters by means other than the application of an alundum coating to the heater wire. It is possible to use insulation formed by ceramic tubes inserted within the cathode (Fig. 11-3b), or to utilize a vacuum separation, where there is a sufficiently great difference between the diameter of the heater and the inside diameter of the cathode (Fig. 11-3c).

In cases in which the utilization of an oxide-coated cathode is

impossible, thoriated carburized cathodes or pure tungsten cathodes are used.

Carburized cathodes, having considerably better emission properties than tungsten cathodes, have gradually replaced the tungsten cathodes. The basic advantages of carburized cathodes are their higher (by a factor of 4-6) working efficiencies and considerably lower working temperatures. There is a considerable decrease in the heating power drawn by high-power transmitting tubes when a carburized cathode replaces a tungsten cathode. Thus, for example, in order to

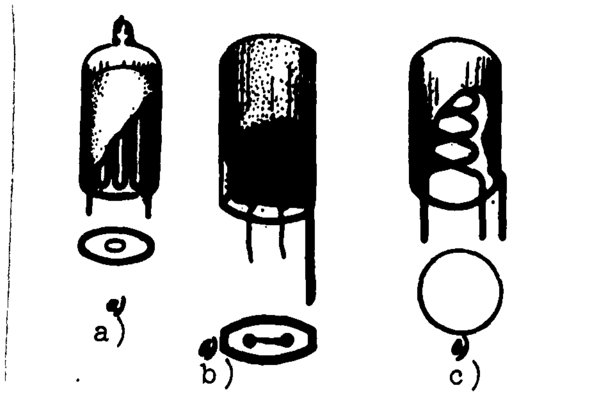

Fig. 11-3. Oxide-coated cathode design for transmitting tubes.

obtain precisely the same output power in a GU-10B tube with a carburized cathode, 525 watts of heating power are needed, while in a GU-89B tube with a pure tungsten cathode, 1380 watts are required, i.e., the heating power with a pure tungsten cathode is greater by a factor of 2.5 in comparison with the heating power consumed by the carburated cathode.

A decrease in heating power for the cathodes of high-power tubes not only yields a direct savings in electric power, which may be very substantial if transmitting equipment is in common use, but also results in a considerable improvement in transmitting-tube operating conditions in that the loads on the electrodes are decreased. A decrease

61-102

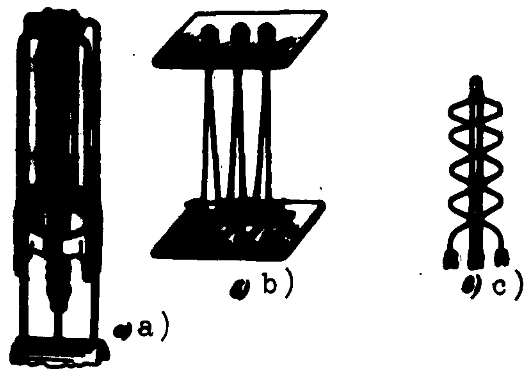

Fig. 11-4. Carburized and tungsten cathode structures for transmitting tubes.

in the power dissipated on the plate and grids of a tube owing to radiation of power from the cathode makes it possible to increase the power evolved on these electrodes owing to load currents, and, consequently, makes it possible to increase the RF power delivered while maintaining the size of the tube unchanged. On the other hand, it is possible to obtain the required values of RF power with a decrease in the size and weight of tubes in comparison with tubes using tungsten cathodes. A decrease in the size and weight of tubes while simultaneously decreasing the heating power required is especially important for tubes utilized in portable transmitting equipment.

Carburized and tungsten cathodes are manufactured as series- or parallel-connected loops and helices (Fig. 11-4). In high-power tubes, the cathodes are sometimes supplied from a three-phase line, in which case the elements of the cathode are wye-connected with a center-tap lead.

A flat loop cathode is utilized in low- and medium-power tubes (up to about 200 watts). In more powerful tubes, either a helix (or double helix) is used or a loop-type cathode with the cathode sections so arranged as to form a cylinder.

61-102

In order to avoid thermal deformation of cathode loops, it is necessary to place them under tension with special springs, as shown in Fig. 11-4. Helical cathodes are normally supported by a center rod.

11-5. TRANSMITTING-TUBE PLATES.

Plate designs are determined by the plate power dissipation, and the method of cooling. In order to determine the necessary radiation surface for plates, we must know, in addition to the power evolved at the plate by the plate current, how much to allow for the power radiated from the cathode and from the grids of the tube so as to reach the plate.

In transmitting tubes with up to 100-150 watts of plate dissipation and oxide-coated cathodes, it is possible to utilize blackened nickel to make the plate; this has a permissible load of up to 1.2 watts per cm^2 of radiation surface. In medium-power tubes with carburized cathodes, the plate materials may be blackened nickel with a permissible load of up to 4.2 watts/cm^2, compressed graphite, molybdenum, and tantalum. Pressed graphite has a very high permissible load — up to 50 watts/cm^2, and is a valuable material for tube plates from the point of view of considerable power with small size. Drawbacks to graphite plates are their high cost and the complexity of production processing, owing to the high brittleness of graphite.

The high-melting metals, molybdenum and tantalum, may be utilized at very high temperatures. The most important advantage of tantalum is its ability to absorb gases at high temperatures. Thus, heated tantalum plates simultaneously serve as getters, which improve the vacuum in a tube. Tantalum is very expensive, which somewhat limits its use.

Examples of plate designs for medium-power tubes are shown in Fig. 11-5.

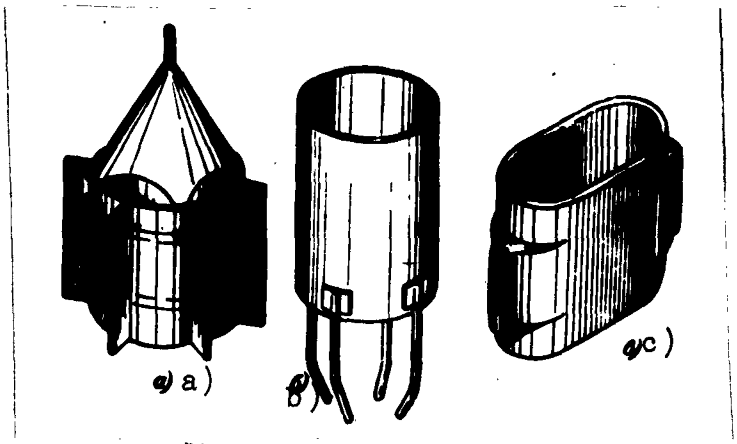

Fig. 11-5. Examples of medium-power tube plate designs.
a) Cylindrical tantalum plate;
b) cylindrical molybdenum plate;
c) graphite plate.

External high-power tube plates are manufactured from red copper. The basic positive qualities of copper, causing it to be utilized in this case, are its high thermal conductivity, imperviousness to gases, which is important for tubes in which the plates form a part of the envelope, and good workability. The working temperature of copper plates should not exceed 400–450°C, since at higher temperatures, intensive evaporation of the copper begins. In practice, the working temperature of copper plates is considerably below that indicated, which is connected with cooling conditions.

When forming a portion of the envelope, copper plates should be sealed into the glass portion of the tube. A copper plate should be sealed into the glass either directly or through a kovar ring. In the first case, at the point of sealing, the end of the plate should be very thin (several hundredths of a millimeter), in order to avoid cracking the seal owing to the sharply different coefficients of expansion of copper and glass. In the second case, the kovar ring should be silver-soldered to the copper plate and the kovar ring in

4-102

turn sealed to the glass portion of the tube. Methods of joining copper plates with glass are shown in Fig. 11-6.

With water cooling of external copper plates, the plates are located in a tank (Fig. 11-7), through which water flows under a pressure of about 3-4 atmos. In cooling the plates, it is necessary to reduce the temperature of the outside surface to the boiling point at the pressure existing in the tank, since upon contact of cold water with the plate surface, a film of steam forms, which impairs the contact of the water with the plate. In addition, when the water boils at the plate surface, scale forms, leading to local overheating of the plate.

In many cases, a water-cooled installation proves to be inconvenient, while it is impossible to use this method in mobile radio stations. Thus, many types of power tubes have plates designed for forced-air cooling. In this case, the outer surface of the plates are provided with fins through which a stream of air is directed from fans (Fig. 11-8). In the majority of cases, the fins of the radiator are soldered to a core or inserted in slots, after which the core of the radiator is soldered to the outside surface of the plate.

Radiator cores are made of copper, while the fins are made of copper or aluminum. Clearly, the amount of power carried off by the radiator at a given temperature depends upon the surface area of the radiator fins and upon the velocity of the air stream. For a given plate length, the surface area of the fins depends upon their number and their width. The number of fins for a given core diameter is limited by the minimum possible fin thickness, and the separation between them at the point where they are connected to the radiator core. A considerable increase in fin width is pointless, since the temperature of sections of fin that are distant from the core proves to be

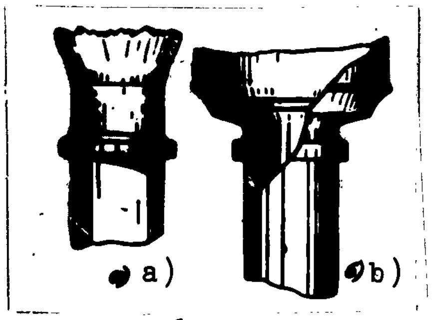

Fig. 11-6. Methods of joining copper plates with a glass envelope.

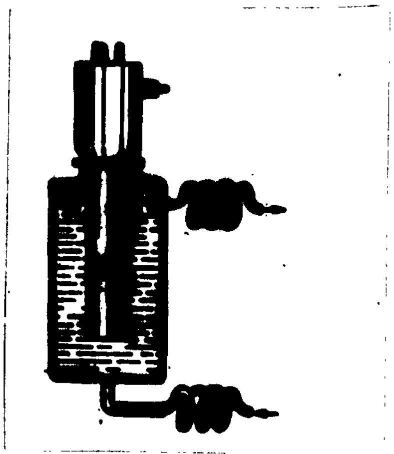

Fig. 11-7. Water cooling of transmitting tube.

considerably below the temperature of the core itself, and the power carried off by the sections proves to be slight. An increase in radiator surface is produced by the utilization of curved fins rather than flat fins (Fig. 11-9).

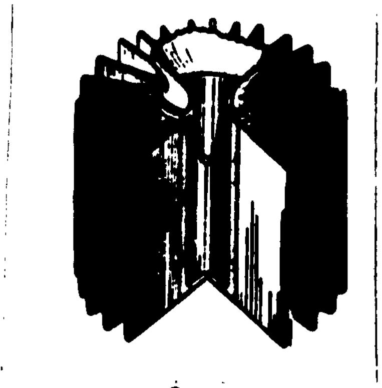

Fig. 11-8. Plate with air-cooling with the aid of a radiator.

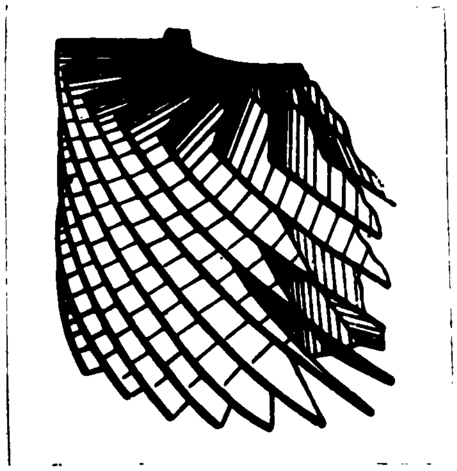

Fig. 11-9. Air-cooling radiator with curved fins.

The quality of the joint between the radiator core and the plate surface has great influence on the amount of power carried off from the plate by an air-cooled installation. Where the joint has poor thermal conductivity, the temperature gradient across the joint

61-10r

increases, which may lead to overheating of the plate. Silver or cadmium solders may be used in soldering the core to the plate; cadmium is more desirable, since cadmium solder has higher heat conductivity than does silver solder and, in addition, the melting point is somewhat higher (321°C), which makes it possible to increase the working temperature of the plate. A silver-copper solder is used to solder the radiator fin to the core.

11-6. TUBE GRIDS.

Grid designs and materials are chosen depending upon the power dissipated on the grids, the type of cathode used in the tube, and the temperature of the remaining electrodes.

In tubes operating at not-too-high power (up to 100-150 watts on the plate) with oxide-coated cathodes, as a rule, molybdenum spiral grids are used, wound on nickel or copper traverses; as in the case of receiving-amplifying tubes, the grid turns are fastened to the traverses either by embedding or by welding. Radiators are widely utilized to cool grids; they are soldered to the traverses. In order to avoid poisoning the cathode with oxides of molybdenum, a protective coating is applied to the grid turns; gold is frequently used for this purpose. At the same time, the gold serves to decrease thermal currents from the grid. In tetrodes and pentodes designed for pulse operation, not only the control grid but the screen grid as well are gold plated, since it is possible for thermal currents to appear from the screen grids to the plate with the tube cut off, which at high plate voltages may lead to the appearance of a very large amount of power on the plate, and to failure of the tube. In tubes with a large average power developed on the grids (for example, in pulsed triodes with oxide-coated cathodes) the grid temperature may reach 600-650°C, which makes gold plating

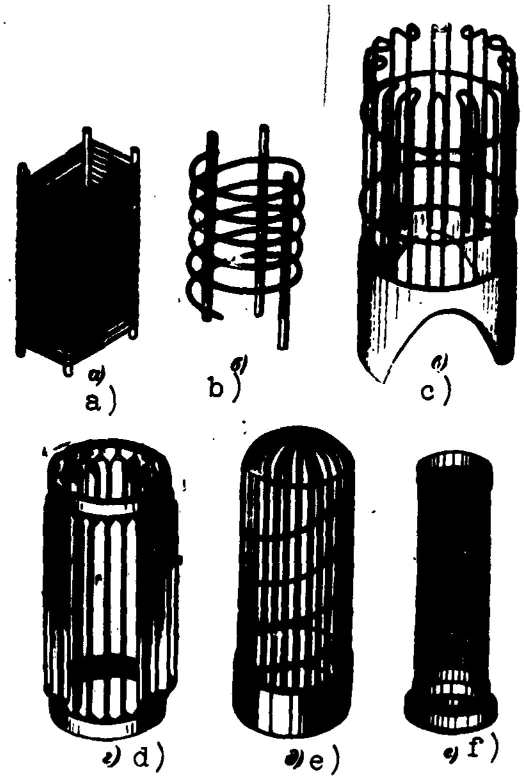

Fig. 11-10. Grid designs for transmitting tubes.

impossible, since at such temperatures, gold easily evaporates in a vacuum. In these cases, platinum-plated grids are used. Good results are also given by a coating of tungsten carbide.

In low-power transmitting tubes, the grids are blackened for cooling purposes; this increases the radiation coefficient of the surface. Such blackening is desirable where plate temperatures are not very high or where plates are of open construction.

For medium-power tubes with tungsten and carburized cathodes, and for high-power tubes, the grids are made of molybdenum wire on molybdenum traverses, and from tantalum. The majority of high-power tubes have cylindrical structures and, consequently, the grids are also made as cylinders. Figure 11-10 shows several grid designs for transmitting tubes. One of the most common is the spiral grid wound on

61-102

traverses, with the turns fastened to the traverses either by welding (Fig. 11-10a) or by means of a wrapping (Fig. 11-10b). In addition to spiral grids, rod-type grids are also used; they are formed of rods parallel to the cathode and fastened into end rings (Fig. 11-10c), or made from lengths of tantalum ribbon, soldered to tantalum rings (Fig. 11-10d). In tubes for pulse operation at medium powers, a combination of a spiral grid with a rod-type grid is encountered (Fig. 11-10e). Such a grid may take the form of a multitraverse spiral grid with a large-pitch spiral and small-diameter traverses.

The surface of molybdenum grids for transmitter tubes is normally zirconium-plated. The zirconium applied to the grid serves, first, as a getter and, second, increases the radiation coefficient of the grid surface. In tubes with tantalum grids, the tantalum itself acts as a getter, with good absorbing qualities appearing at temperatures of from 700 to 1200°C. In order to decrease thermionic emission from tantalum grids and increase the radiation factor from their surface, they are carburized.

Screen grids for medium-power transmitting tubes are sometimes made of a metal fabric woven from molybdenum or tungsten wire, formed into a cylinder and fastened to traverses (Fig. 11-10f).

11-7. ENVELOPES, STEMS, AND INSULATORS.

The sizes and shapes of transmitting-tube envelopes are determined primarily by the amount of power dissipated on the electrodes and radiated through the envelope into space. Envelopes for low-power tubes are normally made from low-melting glasses (lead glasses), supporting a specific load of up to 0.35 watts/cm^2, while in medium- and high-power tubes, high-melting glasses are used (type ZS-5 molybdenum glass and similar borosilicate glasses, the "nonex" type tungsten

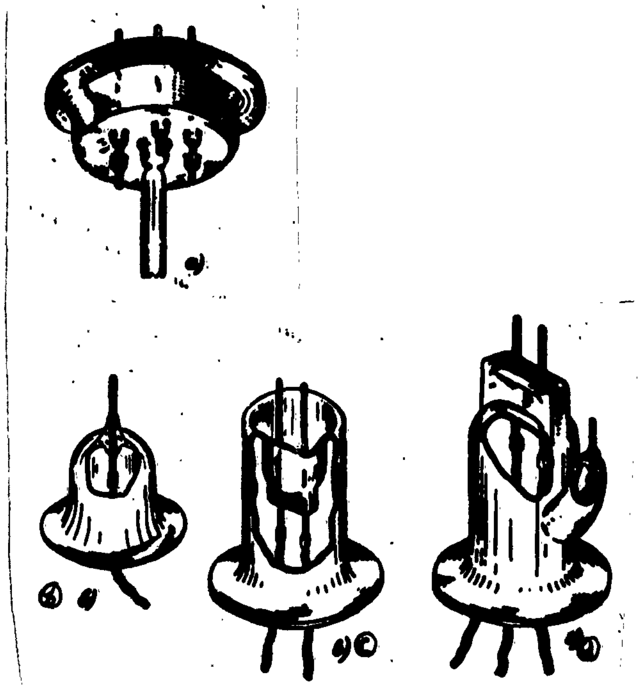

Fig. 11-11. Stem designs for transmitting tubes.
a) Disc stem; b), c) and d) tubular stems.

glass), which permit loads of up to 0.55 watts/cm^2.

In order to decrease the specific load on the glass, the central portion of envelopes may have spherical or barrel shapes. The number of necks for sealing in stems depends upon the type of tube and the arrangement of leads. Tubes are evacuated through an exhaust tip sealed either into the glass of the envelope or into one of the tube stems. The latter method is desirable where a base is utilized that protects the long end of the exhaust tip from mechanical damage.

The glass portion of high-power-tube envelopes bears a high load, since the heavy tube elements are fastened to it. Thus, it is made of high-melting glasses on the order of 2-4 mm thick. The dimensions of this portion of an envelope depend upon the design of the cathode-mounting elements and the grid-mounting elements of the tube.

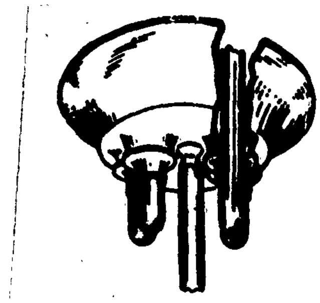

Fig. 11-12. Stem of tube with thick copper leads.

Stem design for transmitter tubes is determined by the magnitude of the working voltages between the electrodes of the tubes, by the electrode currents and the working frequencies. The materials used, to be sealed into the stems of transmitter tubes, are normally molybdenum or tungsten rods with diameters ranging up to 2-3 mm.

Button and disc stems (Fig. 11-11a) are normally utilized in low- and medium-power tubes (up to 100-150 watts on the plate) with inter-electrode voltages not exceeding 600-700 v where the electrodes are connected to leads in the stem. Tubular stems are utilized in all remaining tube designs for medium-power and high-power tubes, with the exception of stems for SHF tubes. In certain types of high-power tubes, it is necessary to make a considerable increase in lead diameter, for example, for the tube cathode leads where the heating current reaches several hundred amperes. This applies to cathode and grid leads for tubes operating at the top of the short-wave band and on VHF. In these cases, fixed copper leads are utilized, soldered to kovar rings which are used to form the seal with the glass (Fig. 11-12).

The electrodes of tubes not operating at high power are mounted in mica and ceramic insulators, which provide for precise mutual location of the electrodes. The utilization of mica insulators is limited by the

fact that mica cannot be heated to temperatures above 500°C. Thus, in the majority of transmitter tubes, the insulators used are made from alundum or steatite ceramic elements. The advantages of ceramic insulators lie in their heat resistance and strength. Ceramic insulators cannot be as precisely manufactured as are mica insulators, but where the interelectrode distances are large, this has no great effect upon parameter variations. In cases in which electrodes must be located with high accuracy, ceramic insulators are augmented with eyelets, brackets, etc.

The possibility of the appearance of surface conductance along ceramic insulators owing to conducting deposits makes them undesirable for utilization where the interelectrode voltages are high. In addition, cracking and fracture of ceramics is possible where there is considerable heating. Thus, in the majority of medium- and high-power tubes, the tube electrodes are mounted directly into glass stems. Light tube elements may be fastened directly to the leads sealed into the stem; in this case, it is possible to seal the lead-support directly into the envelope, without a stem (Fig. 11-13). Larger electrodes are mounted to stems with the aid of special fixtures connected with spring rings and clamps, fastened to the stems (Fig. 11-14).

In mounting tube electrodes on stems, it is necessary to protect points in which leads are sealed into glass or points in which glass is joined to the plates of high-power tubes against radiation from electrodes, particularly from cathodes and plates. Overheating of a joint may lead to cracking of the glass and failure of the tube. In addition, in tubes with high interelectrode voltages, it is necessary to provide protection for joints against high potential gradients, and particularly in the presence of irregularities and sharp ends of electrodes such as, for example, where the plate of a high-power tube is

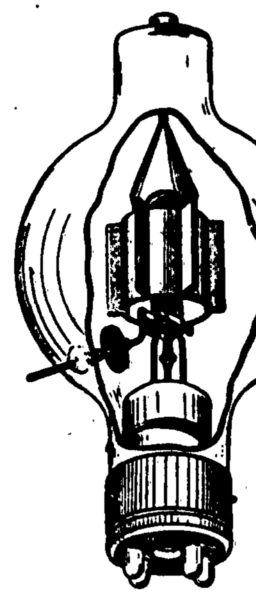

Fig. 11-13. Mounting parts in a medium-power tube.

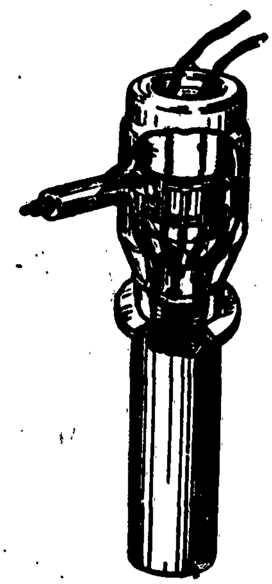

Fig. 11-14. Mounting parts in a high-power tube.

joined to glass. Such protection is accomplished with the aid of special shields which may simultaneously serve as fastening elements.

UNEDITED ROUGH DRAFT TRANSLATION

ELECTRONIC DEVICES (SECOND EDITION)

BY: M. S. Kaufman and G. M. Yankin

English Pages: 653

PART II OF II PARTS (Pages: 394-653)

PREPARED BY:

TRANSLATION SERVICES BRANCH
FOREIGN TECHNOLOGY DIVISION
WP-AFB, OHIO.

Chapter Twelve

ELECTRON TUBES FOR SUPERHIGH FREQUENCIES

12-1. BASIC OPERATING FEATURES OF ELECTRON TUBES AT SUPERHIGH FREQUENCIES.

In modern radio engineering, oscillations at superhigh frequencies are frequently used. In practice, the following superhigh-frequency (SVCh* [SHF]) bands are in use: metric wavelengths for UKV** [VHF] (from 30 to 300 Mc) and three SHF sub-bands — the decimetric (from 300 to 3000 Mc), the centimetric (from 3000 to 30,000 Mc), and the millimetric (above 30,000 Mc). This division is arbitrary, since there are no sharp boundaries between the subdivisions indicated.

In SHF operation, electron-tube properties that were not very noticeable in lower-frequency operation begin to take effect. The possibility of utilizing electron tubes for microwave work is limited, first, by the increase in power losses as frequency rises, resulting in a drop in the useful power delivered by a circuit, second, by the presence of internal tube capacitances and inductances which at microwave frequencies turn out to be of the same order as the capacitances and conductances of external RF circuits and, third, the fact that the electron transit time between electrodes becomes comparable with the period of the generated or amplified alternating voltage.

As an example, let us consider the value of the capacitances and conductances of the intratube leads. The inductance of a wire lead 15 mm long and 0.6 mm in diameter is about 0.18 μh. At a frequency of 3000 Mc, the reactance of this length of wire equals about 3500 ohms. High frequency tube currents create a noticeable voltage drop across the lead

---

* [СВЧ - SVCh - sverkhvysokaya chastota - superhigh frequency.]
** [УКВ - UKV - ul'trakorotkiye volny - ultrashort wave (VHF).]

reactances, and as a result, the tube interelectrode voltages differ sharply from the applied voltages.

It takes very complicated equations to describe the effect of all of these factors upon the operation of tubes in the microwave region. Below we shall consider the basic qualitative relationships determining the structural features of tubes designed for microwave operation.

12-2. TUBE INTERNAL CAPACITANCES AND CONDUCTANCES.

When electron tubes operate in microwave oscillation and amplification circuits, the interelectrode capacitances turn out to be connected in parallel with the external tank circuits (for circuits in which the tuned loops use lumped inductances and capacitances). In order to obtain the necessary gain and power, it is necessary to have a high enough value of tuned-circuit resonant impedance, which depends upon the tuned-circuit constants and the frequency in accordance with the expression, well known in radio engineering:

$$Z_{рез} = \frac{1}{\omega^2 RC},^* \qquad (12\text{-}1)$$

where ω is the angular velocity; R is the resistance, and C is the tuned-circuit capacitance. Since as frequency increases, the resonant impedance of a tuned circuit drops in proportion to the square of the frequency, in order to obtain the required value of resonant impedance, it is necessary to decrease the tuned-circuit capacitance, while selecting an appropriate value of inductance. Even if the capacitance of the tuned circuit itself should be zero, however, the maximum resonant impedance that can be obtained will still be limited by the values

---

* [$Z_{рез}$ - Z_{rez} - $Z_{rezonansnoye}$ - $Z_{resonant}$.]

of tube internal capacitances: the output capacitance for the tube of

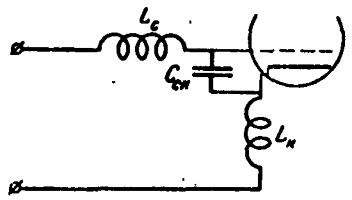

Fig. 12-1. Triode input circuit.

a given stage, and the input capacitance of the tube in the following stage of amplification.

For triodes, the magnitude of the input capacitance under dynamic conditions is determined basically by the grid-plate and grid-cathode capacitances; as we have noted before, the grid-plate capacitance enters into the expression for the dynamic input capacitance with the multiplier $(\mu_d + 1)$. Since both interelectrode capacitances are on the same order, it is the grid-plate capacitance that has the basic effect.

In RF pentodes and tetrodes, the control-grid — plate capacitance has a value that is less by a factor of hundreds than the value for triodes, and its effect upon the dynamic input capacitance of the tube is slight.

The dynamic output capacitance is determined by the tube plate-cathode capacitance. In pentodes, this capacitance is increased owing to the plate-suppressor capacitance, which limits their usefulness at frequencies above 400-500 Mc.

The considerations just mentioned are applicable to tubes operating in common-cathode circuits, in which the input circuit is connected between cathode and grid, and the output circuit between cathode and plate.

Of all of the tube internal inductances involved in common-cathode circuits, the greatest effect is produced by the cathode-lead inductance,

which produces coupling between the input and output circuits of the tube. As a result of this coupling, power is transferred from the grid circuit to the plate circuit; in this case, the amount of power expended by the source of excitation voltage rises in proportion to the square of the frequency. This proves to be especially substantial for SHF transmitting tubes operating class C, i.e., with a large amount of excitation power consumed.

Electrode-lead inductances limit the maximum frequency at which a tube may serve, since these, together with the interelectrode capacitances, form series resonant circuits, connected in parallel with the input and output loops of the circuit. The maximum frequency at which a tube can still operate may not exceed the natural resonant frequency of the loops formed by these inductances and capacitances.

As an example, let us consider a triode input circuit (Fig. 12-1). The inductances of the grid and cathode leads and the grid-cathode capacitance form a tuned circuit resonating at a frequency determined by the well-known formula

$$f_0 = \frac{1}{2\pi\sqrt{LC}}. \qquad (12\text{-}2)$$

In the case under consideration, we shall make the substitution C = $= C_{c.k}$ and $L = L_c + L_k$. If $C_{c.k} = 1.4\,\mu\mu f$, and the total inductance of the cathode and grid leads is $L = 3\cdot 10^{-8}$ henry (these are approximately the values for the best receiving-amplifying triodes with wire leads), then the maximum frequency at which such a triode may operate will equal:

$$f_0 = \frac{1}{2\pi\sqrt{1{,}4\cdot 10^{-12}\cdot 3\cdot 10^{-8}}} = 700\,\text{Mc}.$$

A similar effect is produced by the output capacitance and the plate and cathode lead inductance, which form an output resonant circuit.

These considerations are equally applicable to pentodes.

A substantial effect upon the operation of SHF amplifier circuits is produced by tube transfer capacitance, which determines the coupling between the plate and grid circuits. The resulting feedback may lead to the generation of internal oscillations, sharply increasing the power consumed in the tube, and disturbing the operation of the amplifier.

12-3. ELECTRON TRANSIT TIME. TRANSIT ANGLE.

In studying processes occurring in tubes at moderate frequencies, it may be assumed that the time required for an electron to move between two electrodes will be such that during this period of time the electrical field will remain nearly unchanged, since the period for voltage changes on the electrodes is many times greater than the electron time of flight. In microwave operation, this condition does not hold: the electron transit time between electrodes becomes comparable with the period of the alternating voltage. In a planar diode with a constant plate voltage, let U_a = 25 v, and the plate-cathode spacing be r_a = 0.2 cm. In accordance with formula (5-16)(Chapter Five), let us determine the time required for an electron to pass from the cathode to the plate:

$$\tau_{эл} = \frac{3r_a}{\sqrt{2\frac{e}{m}U}} = 0{,}51\cdot10^{-7}\frac{r_a}{\sqrt{U_a}}[\text{sec}] \approx 0{,}2\cdot10^{-8}\ \text{sec.}\ *$$

If in addition to the constant voltage, an alternating voltage is applied between the plate and cathode of this diode at a frequency of 250 Mc, i.e., with a period $T = 0.4\cdot10^{-8}$ sec, the transit time will turn out to equal half a cycle of this voltage. We should note that this calculation is correct for very small AC amplitudes, that have practically no effect upon the velocity of the electron.

* [$\tau_{эл}$ - τ_{el} - $\tau_{elektron}$ - $\tau_{electron}$.]

The picture of electron motion changes if the alternating-voltage amplitude becomes comparable with the constant voltage, or if the alternating voltage alone is applied between the tube electrodes. In this case, the transit time turns out to be different for electrons leaving the cathode at different instants of time, since they move in the inter-electrode spaces in the presence of various values of electric field strength.

In order to evaluate the influence of electron transit time upon tube operation, we introduce the concept of the transit angle. The transit angle is defined as the angle corresponding to the change in phase of an alternating voltage during the time of flight of an electron from one electrode to another. For example, if the cathode-plate transit time for an electron in a diode equals a quarter cycle of the alternating voltage then, while the electron is passing from the cathode to the plate, the phase of the alternating voltage will change by 90° (or $\pi/2$ [radians]). This means that the transit angle $\theta = 90^{\circ}$. Consequently, the transit angle may be found from the following expression:

$$\theta = \frac{\tau_{pr}}{T} \cdot 360^{\circ}, \qquad (12\text{-}3)$$

where T is the period of the alternating voltage.

In cases where the alternating-voltage amplitude is small, it is possible to find out how the transit angle depends upon the geometry of the interelectrode space and the operating regime of the tube. Substituting into Eq. (12-3) the transit-angle value from expression (5-16), and making the substitution $T = 1/f$, we obtain:

$$\theta = 18 \frac{r_a f}{\sqrt{U_0}} \text{ [degrees]}, \qquad (12\text{-}4)$$

where U_0 is the magnitude of the constant voltage.

For large alternating-voltage amplitudes, where the transit time

and, consequently, the transit angle differ for electrons leaving at different instants of time, the concept of the imaginary transit angle is introduced. The imaginary transit angle is defined as the transit angle computed for a constant voltage equal to the amplitude of the alternating voltage. This angle may be calculated in accordance with formula (12-4), if we substitute into it in place of the constant voltage U_0 the amplitude of the alternating voltage U_1.

Formula (12-4) holds if a tube is operated under space-charge-limited current conditions. In practice, it is frequently necessary to deal with an operating regime in which the effect of the space charge is negligible (for example, the motion of electrons in the grid-plate space of a triode). In this case, the transit angle proves to be lower, and it may be computed from the formula

$$\theta = 12\frac{r_a f}{\sqrt{U_0}} \text{ [degrees]} \qquad (12\text{-}5)$$

The effect of electron transit time on tube operation may be neglected where the transit angle has a value $\theta < 0.1\pi$, i.e., where the transit time is less than one-twentieth of a cycle of the alternating voltage. At larger angles θ, the transit time has a noticeable effect upon tube operation.

12-4. EFFECT OF ELECTRON TRANSIT TIME UPON TUBE OPERATION.

For large transit angles, the magnitude of tube electron currents may not be determined solely in terms of the number of electrons passing over to an electrode in unit time. This is connected with the fact that electrons moving in interelectrode spaces induce currents in electrodes, currents whose magnitudes depend upon the number of electrons in motion and upon their velocities. As a result, a substantial current may appear even in an electrode upon which no electrodes whatsoever are impinging.

Let us consider this in more detail, using the example of electrons moving in a plane diode (Fig. 12-2). An electron lying between cathode and plate induces on both electrodes positive charges whose sum, in absolute value, equals the charge on the electron. If the electron is near the cathode, then the charge induced on the cathode exceeds the charge induced on the plate. As the electron moves, the magnitude of the charges induced upon each electrode varies, which is equivalent to the appearance of an electrode current.

The magnitude of a current induced upon an electrode is determined by the rate of change of the magnitude of the induced charge, i.e., the rate of translation of the electron along the cathode-anode direction. Where there is an external circuit connecting the cathode to the plate, the change in the charges upon these electrodes causes an equalizing current to appear in the circuit.

If we neglect the effect of space charge, in a diode with a constant voltage upon the plate, the distribution of potential between plate and cathode may be assumed to be linear. In such a field, an electron will move with constant acceleration, and the current induced on the plate throughout the entire plate-cathode transit time will rise linearly. Consequently, each electron leaving the cathode creates a plate-current pulse equal in length to the transit time (Fig. 12-3). The total plate current is found in terms of the sum of the current pulses induced by all electrons moving in the plate-cathode space.

Let us consider still two more cases of formation of induced currents in a diode in the presence of an alternating voltage between plate and cathode. In the first case, we will assume that the amplitude of the alternating voltage is small in comparison with the constant voltage and, consequently, that the transit angle for all electrons is identical. The stream of electrons leaving the cathode is mentally

divided into separate layers, with the electron density in each layer

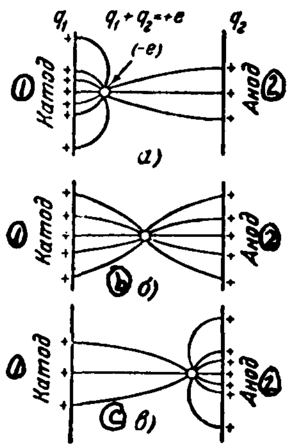

Fig. 12-2. Formation of an induced current in a diode.
1) Cathode; 2) plate.

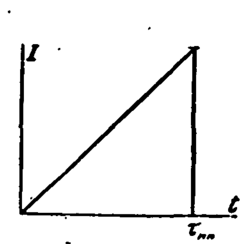

Fig. 12-3. Current pulse induced by a single electron.

dependent upon the phase of the alternating voltage at the instant the layer is defined at the cathode. Maximum density will evidently occur in that layer which leaves the cathode when the alternating plate voltage is at a maximum. Upon moving toward the plate, each of these layers sets up a plate-current pulse whose amplitude depends upon the number of electrons in the layer. The resultant plate current will contain an alternating component which will lag behind the phase of the alternating voltage by the transit angle. If the effect of the alternating voltage upon the electron transit angle cannot be neglected, the phase shift will differ from the computed imaginary transit angle.

For our second example, we will examine the case in which only an alternating voltage at a microwave frequency is applied to the plate. Electrons leaving the cathode during the positive half cycle move toward the plate. A portion of these electrons, in moving toward the plate, fall into a retarding field, since during the transit time, the plate can become negative. Some electrons reach the plate at lower velocities, while the rest will return to the cathode. The larger the

transit angle, the smaller is the number of electrons that can reach the plate. It has been computed theoretically, for example, that with an imaginary transit angle of 120°, about 50 per cent of the electrons will return to the cathode. The current induced on the plate will have an alternating component at the fundamental frequency, shifted in phase relative to the alternating voltage. As the frequency goes up, it is possible to reach such values of transit angle that no electrons at all will strike the plate, and the current through the diode will be purely capacitive in nature.

Thus, an increase in the transit angle leads to the appearance of a phase shift between the current on the electrodes and the controlling alternating voltage.

The effect of electron transit time upon triode operation is primarily felt in the appearance of a conductance component in the grid-cathode circuit, which leads to an increase in the power consumed in the excitation circuit. When electrons move in the cathode-grid space and in the grid-plate space, an induced current appears upon the grid. The electrons moving from grid to plate induce a current whose direction is opposite to the direction of the current induced by electrons moving from cathode to grid. At small transit angles, the magnitude of the current on both sides of the grid is the same, and the induced currents cancel. With large transit angles, the current in the grid-cathode space will differ from the current in the grid-plate space, and an alternating current will be induced on the grid that is shifted in phase with respect to the alternating voltage of the grid. The active component of this current will also determine the power consumed in the grid circuit.

Since there is no electronic current on the grid (with a sufficiently large negative bias), the power consumed in the grid circuit cannot

be evolved directly at the grid. A physical interpretation of this power consumption may be given if we consider the interaction of the electrons with the high-frequency field of the grid. A portion of the electrons leaving the cathode during the positive half cycle of the alternating grid voltage will be accelerated along the path from cathode to grid by the high-frequency field, gaining energy in the process. If after the electrons reach the plane of the grid, the voltage becomes negative, the field set up by this voltage again turns out to be an accelerating field in respect to the electrons and, consequently, the electrons will gain energy from the high-frequency grid field along the entire path from cathode to plate. The kinetic energy acquired by these electrons is given up at the plate, and appears as heat. Similarly, a portion of the electrons leaving the cathode during the negative half cycle of the alternating grid voltage will move in a retarding high-frequency field along the entire path to the plate, and will give up their energy to the grid circuit. Since more electrons will leave the cathode during a positive half cycle than during a negative half cycle, the energy taken from the grid circuit turns out to be greater than that returned to it. As a result, power is drawn from the source of grid excitation, and the power reappears at the plate as heat.

The greater the electron transit angle, the more electrons fall into the "reverse" phase of the alternating grid voltage after reaching the grid and, consequently, the greater the power losses of the excitation source. A mathematical analysis of this phenomenon shows that the power loss in a grid circuit rises in proportion to the square of the frequency.

The size of the transit angle in the cathode-grid space has a great effect upon triode operation. As this angle increases, at sufficiently large alternating-voltage amplitudes, a portion of the

electrons leaving the cathode cannot pass beyond the grid, and are returned to the cathode. As a result, the number of electrons experiencing high-frequency oscillations drops and the RF power also drops. In addition, some of the electrons retarded in the grid-cathode space do not succeed in reaching the cathode, and are again turned back to the grid and pass into the grid-plate space. These electrons may fall into a phase that is "neutral" with respect to the alternating voltage on the plate. In other words, these electrons may induce current pulses on the plate that will lead to damping of oscillations.

Both theory and practice of designing microwave tubes have shown that the maximum transit angle in the cathode-grid space at which the tube will still operate is an angle close to 120°.

An increase in the transit angle in the grid-plate space increases the phase shift between the alternating components of the plate current and the plate voltage, resulting in a drop in RF power.

As may be seen from formulas (12-4) and (12-5), there are two ways to decrease the transit angle: a decrease in interelectrode spacing and an increase in working voltages. The first is the basic method for decreasing transit angles. There are limitations upon decreasing grid-cathode spacing from the point of view of the production technology for these elements and accuracy of tube assembly. As has been shown, when the grid-cathode spacing is decreased, it is simultaneously necessary to decrease the grid pitch and, consequently, the grid-wire diameter. In modern microwave tubes, this spacing ranges from 15 to 200 microns. It should be noted that bringing the grid closer to the cathode decreases the transit angle considerably more rapidly than it would appear from the formulas given. This is explained by the fact that in modern tubes with cathodes that have high emission, the transit time is determined not by the grid-cathode spacing but by the separation

of the grid from the space charge that forms the virtual cathode.

An increase in the amplitude of the working voltages leads to a sharp rise in the thermal loads on electrodes, and the cathode mean-current load. The resulting power losses are not compensated by a gain in RF power owing to a decrease in transit angle, while cathode life decreases sharply. Thus, as a rule, microwave tubes are operated under lower working voltages than are long-wave tubes.

12-5. ENERGY LOSSES IN MICROWAVE OPERATION OF TUBES.

Energy losses in tubes operating in the microwave region may be classified as follows: 1) Losses in electrode leads where the alternating RF voltages are applied, and on the electrodes themselves; 2) losses owing to radiation of energy; 3) dielectric losses in insulators; 4) energy losses owing to RF currents induced in tube elements to which RF alternating voltages are not applied. Let us consider the reasons for these losses.

Energy losses in electrodes and leads occur owing to skin effect in the microwave region. As a result, there is a sharp rise in lead resistance, since current is conducted only by a thin surface layer. The theory of the electric field in conductors gives the following approximate expression for the ratio of the AC resistance of a conductor to the DC resistance:

$$\frac{R}{R_{\text{пост}}}=\frac{d_{\text{вв}}}{4}\sqrt{\pi f \mu\mu_0\gamma}+0,25.\text{*} \qquad (12\text{-}6)$$

In this formula, d_{vv} is the lead diameter; f is the AC frequency; μ and μ_0 are respectively the relative magnetic permeability of the

* [$R_{\text{пост}}$ - R_{post} - $R_{postoyannaya}$ - R_{direct}.
$d_{\text{вв}}$ - d_{vv} - d_{vvod} - d_{lead}.]

lead material and the magnetic permeability of free space; γ is the conductivity of the lead material. All quantities are expressed in the absolute practical system (MKSA).

It is clear that the increase in lead resistance is greater where materials with high magnetic permeability are used (iron, nickel, and many other alloys). Thus, the resistance of a round iron wire 1.0 mm in diameter increases by nearly a factor of 20 at a frequency of 3000 Mc, while the resistance of a nickel wire of the same dimensions increases by about six times in comparison with the DC resistance. Thus, in tubes for the decimetric and centimetric bands, nonmagnetic materials are used for the leads. One way to decrease lead losses is to copper or silver the lead surfaces, which decreases the surface resistance. Here it is necessary to pay special attention to the tightness of the surface layer, since the normally utilized electrolytic method of plating produces a porous metal structure. It is necessary to protect the lead surfaces against oxidation, which sharply increases lead resistance.

The same considerations also apply to the materials of the tube electrodes themselves; the RF currents also flow along the surfaces of the electrodes. This is the reason that it is desirable to manufacture internal elements and electrodes of tubes from materials with good conducting properties (for example, copper) and to silver-plate their surfaces. Electrodes, too, cannot be allowed to oxidize or to accumulate various films of particles on their surfaces that have poor conductance.

Electrode and lead losses increase in microwave operation, since there are large capacitive currents flowing at values determined by the interelectrode capacitances, the amplitudes of alternating voltages, and the frequency. These losses increase with frequency, first, owing to an increase in the capacitive interelectrode conductance, and, second, owing to an increase in the resistance of the current-carrying

portions of the tube.

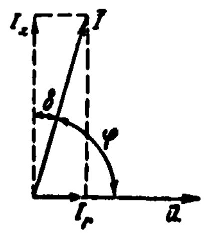

Fig. 12-4. Determination of dielectric loss angle.

A substantial amount of microwave energy may be radiated into space from the current-carrying section of a tube and circuit. In the microwave region, the length of current-carrying conductors turns out to be on the same order as the wavelength, and thus the radiation of energy into space from these conductors becomes considerable. As we know from radio engineering, there is a considerable drop in radiation losses when coaxial two-conductor lines are used consisting of two concentric cylindrical conductors. Thus, tubes designed for microwave operation and in particular tubes for the decimetric and centimetric ranges should provide for connection to coaxial lines.

Large power losses may also be caused by induction currents in tube electrode circuits that are not directly connected with RF loops in the circuit. The heater circuit is an example of this. In order to decrease such losses, it is necessary to carry out careful shielding of auxiliary elements against the effect of the RF fields.

In the decimetric and centimetric ranges, RF losses in insulators become appreciable. These losses are associated with processes occurring during the periodic change in dielectric polarization; the magnitude of the losses depends upon the properties of the dielectric and the frequency of the alternating voltage. Dielectric properties are normally characterized by the so-called loss angle δ. The concept of the loss

angle is entailed by a consideration of processes taking place in a capacitor. The presence of dielectric losses in a capacitor leads to the appearance of an in-phase current component through the capacitor. Thus, the current through a capacitor leads the voltage not by 90°, but by a somewhat smaller angle, as shown in the vector diagram (Fig. 12-4).

The angle formed by the difference between 90° and the lead angle φ is called the loss angle, and characterizes the properties of the dielectric. Clearly, the greater the loss angle, the greater the in-phase current component, and the greater the thermal losses in the insulation.

Designating the pure resistance of a capacitor by r and the reactance by $X = 1/\omega C$, we may find the magnitude of the in-phase current component through a capacitor I_r and the reactive component I_x:

$$I_r = \frac{U}{r}; \qquad (12\text{-}7)$$

$$I_x = \frac{U}{X} = U\omega C. \qquad (12\text{-}8)$$

From the vector diagram

$$\operatorname{tg}\delta = \frac{I_r}{I_x} = \frac{1}{\omega C r},* \qquad (12\text{-}9)$$

from which we determine the pure resistance

$$r = \frac{1}{\omega C \operatorname{tg}\delta}. \qquad (12\text{-}10)$$

Consequently, the insulation resistance drops and the dielectric losses rise as the alternating-voltage frequency goes up. Thus, in tubes for microwave operation, it is necessary to utilize materials

* [tg - tan.]

with low loss angles for envelopes and tube internal insulators.

Table 12-1 shows that the best material from the point of view of dielectric losses are mica and special ceramics, as well as certain types of glass.

It is also clear from the table that for RF tube leads, it is necessary to utilize carbolite bases.

TABLE 12-1.

1) Материал	2) Тангенс угла потерь
3) Карболит (для цоколей)	$(400+800)\cdot 10^{-4}$
4) Свинцовое стекло	$(8\div 12)\cdot 10^{-4}$
5) Стекло „нонекс"	$(4\div 6)\cdot 10^{-4}$
6) Стеатитовая керамика (для баллонов)	$(10\div 14)\cdot 10^{-4}$
7) Алундовая керамика (для баллонов)	$(3\div 4,5)\cdot 10^{-4}$
8) Слюда мусковит	$(1,5\div 2)\cdot 10^{-4}$

1) Material; 2) loss tangent; 3) carbolite (for bases); 4) lead glass; 5) "nonex" glass; 6) steatite ceramic (for envelopes); 7) alundum ceramic (for envelopes); 8) common mica.

12-6. AMPLIFIER TUBES FOR THE METRIC (VHF) BAND.

We may draw the following conclusions from a consideration of the causes for deterioration in the operation of electronic tubes in the microwave region: 1) Tube designs for microwave operation should provide for the maximum decrease in interelectrode capacitances and lead inductances; 2) the interelectrode spacings and working voltages on the electrodes should provide the minimum electron transit angles; 3) in order to provide adequate circuit efficiency, it is necessary to take steps to decrease RF losses in structural elements and tube and circuit elements.

For tubes operating in VHF bands, the basic problem is to decrease interelectrode capacitances and lead inductances. In receiving-amplifying tubes, the problem is solved by decreasing the dimensions of the tube electrodes, and by decreasing lead length while simultaneously separat-

ing them from each other. Miniature and subminiature tubes are examples of such designs. We should note that miniature-tube design has as its goal not only an improvement in the RF properties of the tubes, but also a decrease in their size, weight, and power consumption, which in turn leads to a decrease in the size and weight of radio equipment. Thus, other types of miniature and subminiature tubes are designed, intended for service not only at high frequencies, but for audio-frequency-amplification service as well.

A decrease in electrode surface area leads, first, to a need for a substantial decrease in interelectrode spacings in order to provide the required tube characteristics. This refers primarily to the cathode-control grid spacing which determines the tube transconductance. In several types of tubes, this spacing may be on the order of 40-50 microns, which technically complicates the manufacture of such tubes owing to the need for making tube elements with very small tolerances.

A second consequence of a decrease in the electrode dimensions is a decrease in the permissible power dissipation on the plate, grids, and envelope of the tube. This in turn limits the possibility of obtaining high enough RF power in amplifier circuits. In the majority of miniature and subminiature tubes, the plates and envelopes operate with the maximum permissible values of specific power dissipation.

In order to decrease capacitances between leads and lead inductances, tubes are made without bases and with short leads. Miniature (button-type) tubes are manufactured with button stems; in this case, the leads simultaneously act as pins. The plate and control-grid leads are separated as far as possible from each other (Fig. 12-5a). In subminiature tubes having stems formed by stamping the end of the envelope, the leads are located in a single plane in such manner that the plate and control-grid leads are shielded from each other by the leads of the

other electrodes (Fig. 12-5b).

As we have already said, the inductance of the cathode lead has a substantial effect upon the magnitude of the tube input resistance; this lead is a common element in the plate and grid circuits. In order to decrease the effect of cathode-lead inductance, tubes for the metric and decimetric ranges are constructed with two cathode leads, which makes it possible to separate the plate and grid circuits (Fig. 12-6). An example of a tube constructed in this manner is the 6Zh1P [6AK5]

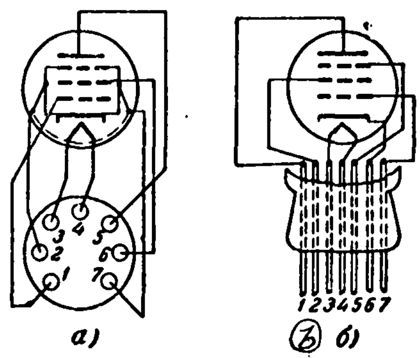

Fig. 12-5. Base arrangement for miniature (a) and subminiature (b) pentodes.

miniature pentode, whose base arrangement is shown in Fig. 12-5a.

A second type of electronic-tube design for the metric range is the all-glass construction. Such tubes use a button stem having a large circumference along which the leads are located (Fig. 12-7).

The types of tube described find wide application at frequencies up to 150-200 Mc. At somewhat higher frequencies (wavelengths of 60-70 cm), it is possible to use "acorn" type tubes. The envelopes of such tubes consist of two bowl-shaped halves sealed together by a so-called welt-type seal. The tube electrode leads are located along the periphery of the seal, which considerably decreases the mutual capacitances. In acorn pentodes, the cathode, heater, and screen and suppressor leads are sealed into the welt seam. The plate and control-

grid leads are located in necks in both sections of the envelope (Fig. 12-8). The interelectrode capacitances and interelectrode spacings in "acorn" tubes are nearly the same as the corresponding quantities in high-frequency miniature and subminiature tubes. Their

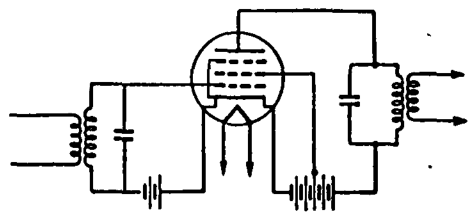

Fig. 12-6. Circuit for connecting a pentode having separate cathode leads.

ability to operate at higher frequencies is occasioned by the location of the leads which is more convenient for connection into the circuit, and which decreases the capacitance between current-conducting wires in the circuit and decreases their inductance.

Triodes operating in the common-grid (grounded-grid) circuit suggested by M.A. Bonch-Bruyevich have found wide application for oscillation and amplification in the microwave region. In this circuit (Fig. 12-9), the output tuned circuit is connected between the grid and the plate. In this case the output capacitance is the grid-plate capacitance, and the transfer capacitance is the plate-cathode capacitance. By decreasing the penetrance of the grid, it is possible to make the transfer capacitance so small that stable operation of the circuit at high frequencies is provided.

As a rule, miniature and subminiature receiving-amplifying triodes are used in grounded-grid circuits. Since in such a circuit the grid lead is common to the input and output circuits, in order to decrease lead inductance and to isolate the circuits, the triode grids are furnished with not one but several leads. The triodes described

may operate in grounded-grid circuits at frequencies up to 500-600 Mc.

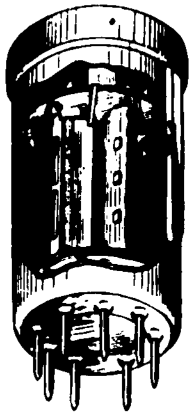

Fig. 12-7. Glass VHF beam pentode.

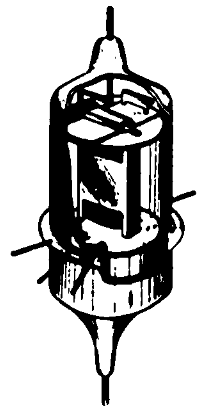

Fig. 12-8. "Acorn" pentode.

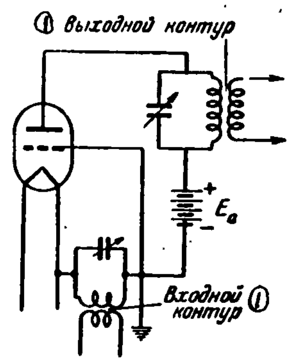

Fig. 12-9. Grounded grid amplifier circuit.
1) Output circuit;

12-7. TRANSMITTING TUBES FOR THE METRIC BAND.

In designing VHF transmitting tubes, the basic problem, as in the case of receiving-amplifying tubes, is a decrease in interelectrode

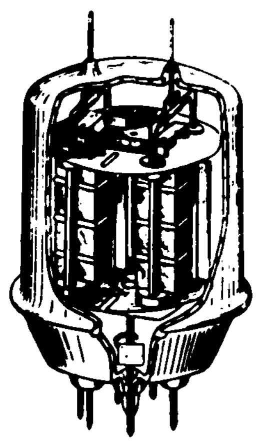

Fig. 12-10. Dual VHF beam-power transmitting tetrode.

capacitances, lead inductances, and electron transit time. The large amounts of power dissipated on transmitting-tube electrodes do not

61-107

permit any substantial decrease in electrode size. On the other hand, the high interelectrode voltages decrease electron transit time, which makes it possible to make the interelectrode spacings so large that the capacitances between the electrodes themselves are decreased. Thus, in designing VHF transmitting tubes, the tube internal capacitances are reduced by shortening the leads and locating them rationally, which should also provide for connecting the tube into the circuit in such manner that the capacitances between current-carrying wires are minimized.

The large alternating currents carried by transmitting-tube electrodes require that lead inductances be as low as possible.

Tetrodes and pentodes with continuous plate dissipations on the order of 500 watts are normally utilized as VHF transmitting tubes; in many cases, two tetrodes or pentodes are combined into a single envelope. Such tubes use glass envelopes with button stems and no bases (Fig. 12-10). The plate leads are brought out through the top of the envelopes in order to decrease the plate – control-grid capacitance. In order to decrease the power dissipated on the screen grid, the majority of tubes utilize a beam-shaping arrangement of the control and screen grids. Depending upon tube power, oxide-coated or thoriated carburized cathodes are used.

A further rise in operating frequency is associated with a sharp drop in the useful RF power delivered by transmitting tetrodes and pentodes. When the tubes are operated in oscillator circuits with constant excitation, an instability appears that is caused by an increase in the transfer conductance owing to the increasing influence of screen-grid-lead inductance. Thus, at frequencies above 200-250 Mc, pentodes and tetrodes lose their advantages over triodes.

VHF transmitting tubes designed for continuous plate dissipations

above 1 kw are in the majority of cases made as triodes operating in grounded-grid circuits. The tubes are designed with copper plates and forced air or water cooling is used.

As we have already said, a decrease in inductance is possible as

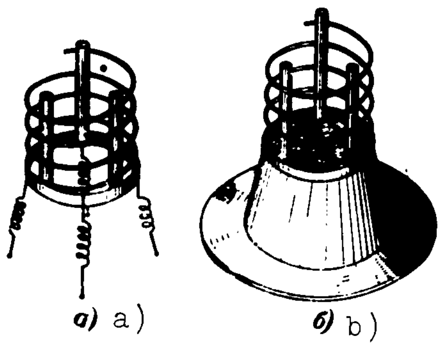

Fig. 12-11. Transition from several leads to a disc-type grid lead.

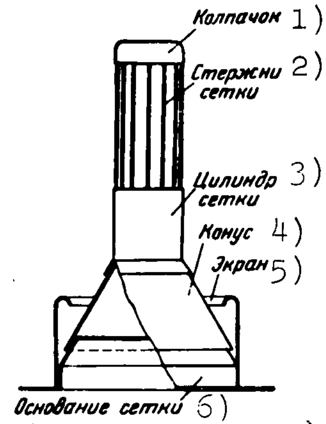

Fig. 12-12. Design of a grid for a high-power UHF transmitting triode.
1) Cap; 2) grid rods;
3) grid cylinder; 4) cone;
5) shield; 6) grid base.

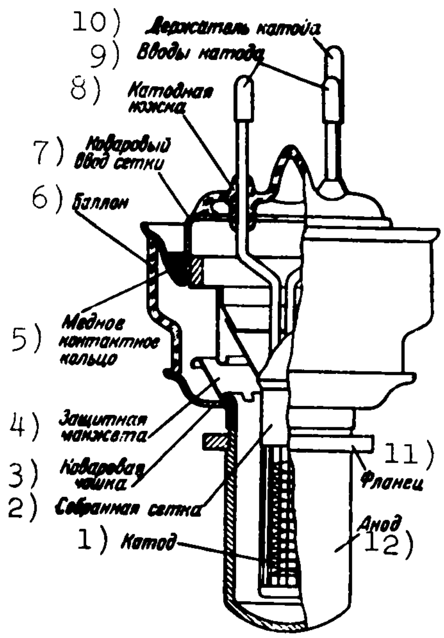

Fig. 12-13. High-power VHF water-cooled transmitting triode.
1) Cathode; 2) assembled grid;
3) kovar ring; 4) protective skirt;
5) copper contact ring; 6) envelope;
7) kovar grid lead; 8) cathode seal;
9) cathode lead; 10) cathode support;
11) flange; 12) plate.

a result of increasing the number of parallel leads to a single electrode, since the equivalent inductive reactance of several inductive reactances connected in parallel is less than the reactance of each of them (Fig. 12-11a). The limiting case for increasing the number of leads is the utilization of a disc-type lead which, for all practical purposes, can be assumed to be inductionless (Fig. 12-11b). In the majority of high-power transmitting tubes for the metric band, the grids use disc-type leads (Fig. 12-12). In such a design, a rod-type or wound grid is fastened with the aid of a cone to a nickel base which in turn is bolted to a kovar disc that acts as a lead. Where the kovar ring is bent, a copper ring is soldered to it, and the wires connecting the grid of the tube into the circuit are run to it. The great advantage of this type of lead is the possibility of making a circuit connection with the aid of a tube having low inductance. On one side of a kovar lead, the tube cathode seal is soldered, and on the other, a glass cylinder separating the grid input from the plate (Fig. 12-13).

Transmitting tubes of this design may deliver up to 250 kw of RF power at a working frequency of up to 120 Mc. At higher frequencies, the RF power drops, owing basically to an increase in the transit angle in the grid-plate space. The decisive influence of the transit time in this space is caused by the fact that in grounded-grid circuits, the output tuned circuit is connected between plate and grid, and an increase in the transit angle leads to an increase in the phase shift between the tuned-circuit voltage and the alternating plate current, and, consequently, to a drop in the RF power in the tuned circuit, which equals:

$$P_{\kappa} = I_{a.\text{эфф}} U_{\kappa.\text{эфф}} \cos\varphi,^{*} \qquad (12\text{-}11)$$

* [эфф - eff - effectivnyy - effective.]

where ψ is the angle defining the phase shift.

A drop in the transit time resulting from moving the plate and grid closer together does not yield positive results, since in this case, the grid-plate capacitance increases, which decreases the resonance impedence of the tuned circuit and, consequently, the RF power developed.

12-8. MICROWAVE TANK CIRCUITS.

At decimetric and centimetric wavelengths, the capacitances and inductances of the conductors connecting a tube with the tank circuit have a very substantial effect. Thus, a characteristic of microwave electron-tube design is the possibility of connecting the tube electrodes directly to the tank circuit. In microwave oscillators and amplifiers, the electron tube is an organic part of the tuned circuit and is connected to it with no intermediate conductors.

The most common types of microwave tank circuits are short-circuited two-conductor lines and coaxial lines. The simplest tank circuit is the two-conductor line, which takes the form of two parallel wires short circuited by a jumper (Fig. 12-14). The inductance of such

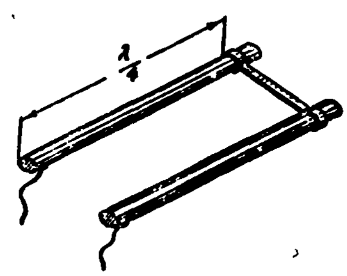

Fig. 12-14. Tank circuit formed of two-conductor line.

a line is determined by the inductance of the conductors, and the capacitance by the capacitances between the individual line elements. Such a line is a circuit with distributed constants. From the theory

of electric circuits with distributed constants, it is known that a short-circuited two-conductor line represents a capacitive reactance with respect to alternating current for a length of line L lying between a quarter wavelength and a half wavelength, and an inductance reactance where the line is more than a half wavelength but less than three-quarters of a wavelength. With a line length equal to an odd number of quarter wavelengths ($\lambda/4$, $3\lambda/4$, $5\lambda/4$, etc.), resonance occurs, and the line appears as a high resistance to alternating current of the appropriate frequency. The value of the resonant impedance of such a line depends upon the inductance of the conductors themselves, their separation, and the resistance of the conductors per unit length.

A drawback to the use of a short-circuited parallel-conductor line as a tank circuit are the considerable amounts of power lost in it owing, first, to the large resistance of the conductors and, second, owing to the radiation of power into the surrounding space. A coaxial line does not suffer from the disadvantages; it takes the form of two coaxial cylinders, short-circuited at one end. A coaxial line may be represented as a product of the revolution of a two-conductor short-circuited line, as shown in Fig. 12-15. As in the case of a parallel line, coaxial-line resonance occurs with a line length equal to an odd multiple of a quarter wavelength. Since at very high frequencies an electric current is propogated along the surface of the conductors, in order to decreases losses owing to resistance, it is necessary to increase the conducting surface. Evidently, the surface of the cylinders in a coaxial line is considerably larger than the surface of the conductors in a similar parallel-conductor line and, consequently, losses owing to the resistance of a coaxial line are considerably lower than the losses in a parallel-conductor line. There are almost no losses owing to radiation of power from a coaxial line, since the

electromagnetic oscillations turn out to be confined to the space between the cylinders of the line.

If a lumped capacitance is connected to a two-conductor line or

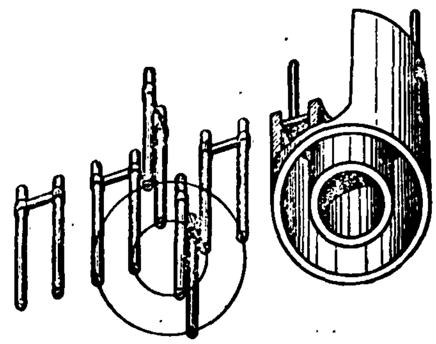

Fig. 12-15. Transition from two-conductor line to coaxial structure.

to a coaxial line, the length of the line corresponding to resonance turns out to be less than a quarter wavelength. The larger the lumped capacitance, the shorter the length of the line. Thus, coaxial lines normally have sliding shorting discs, which are moved in order to tune the line to resonance.

Still higher frequencies may be obtained in the cavity resonator suggested by M.S. Neyman. Let us examine the transition from an ordinary tank circuit with lumped constants to a hollow (cavity) resonator. The natural resonant frequency f_0 of each of the two parallel-connected circuits (Fig. 12-16a) depends upon the value of the inductance L_1 and capacitance C_1. Decreasing the inductance by reducing the number of turns in the coil, and decreasing the capacitance, we will obtain a circuit with minimum values of inductance and capacitance and, consequently, with the maximum resonant frequency (Fig. 12-16c). If such a "limiting" tank circuit is rotated about its own vertical axis, it will form a closed toroidal surface, as shown in Fig. 12-17 (a — vertical

section; b — horizontal section). When specific conditions obtain within such a surface, it is possible to excite electromagnetic oscillations

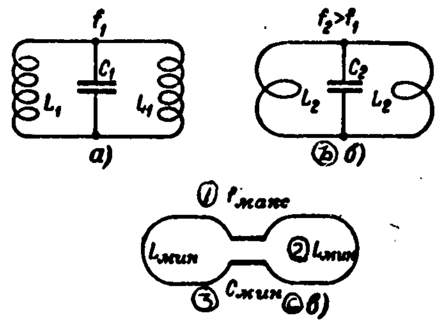

Fig. 12-16. Transition from tank circuit with lumped constants to cavity resonator.
1) f_{max}; 2) L_{min}; 3) C_{min}.

and, consequently, the cavity resonator may play the role of a tank circuit. The lines of force of the electric and magnetic fields are contained within the resonator, since the metal walls shield the cavity from the space outside. Consequently, the cavity resonator has the same advantages as the coaxial line — absence of radiation to the surrounding space, and a large surface which provides a low line resistance.

Depending upon the linear dimensions and shape of resonators,

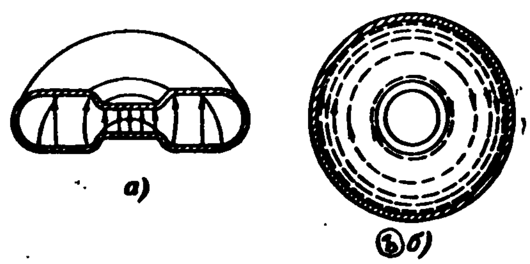

Fig. 12-17. Distribution of lines of force of electrical (a) and magnetic (b) fields in a cavity resonator.

it is possible to excite within them oscillations having a large

number of resonant frequencies, each of which corresponds to a specific configuration of the electromagnetic field. In order to excite undamped oscillations, it is necessary to supply energy to the resonator

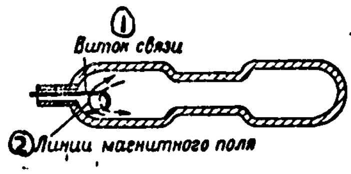

Fig. 12-18. Coupling loop in cavity resonator.
1) Coupling loop;
2) magnetic-field lines.

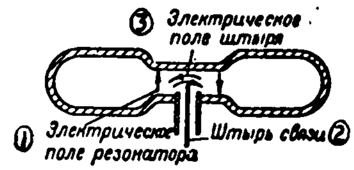

Fig. 12-19. Capacitive coupling to cavity resonator.
1) Resonator electric field; 2) coupling probe; 3) electric field of probe.

from an external source. Energy is transmitted to a resonator by either magnetic (inductive) or capacitive coupling with the source of oscillatory energy. Where magnetic coupling is used, a coupling loop is introduced into the resonator cavity (Fig. 12-18), and the end of the loop is connected to a wall of the resonator. For the tightest coupling, the loop is so arranged that its plane will be normal to the direction of the resonator internal magnetic field. Capacitive coupling is carried out by introducing a coupling probe into the resonator cavity (Fig. 12-19). Coupling loops and probes may be utilized to remove RF energy from resonators, if oscillations are excited in resonator cavities by some other method.

12-9. ELECTRON TUBES FOR DECIMETRIC AND CENTIMETRIC WAVELENGTHS.

Electron-tube design for the decimetric and centimetric bands provides for the possibility of connecting the tube electrodes directly to the cylinders of a coaxial line or with cavity resonators.

One of the common types of microwave triodes is the lighthouse-tube triode (Fig. 12-20), which has flat electrodes with disc seals.

The plate of the tube is made in the form of a cylindrical rod whose lower end, facing the cathode, serves as the working surface. To the

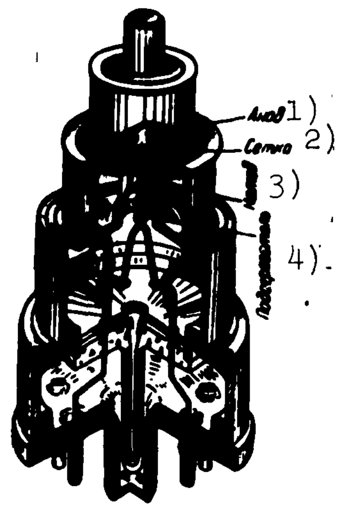

Fig. 12-20. Lighthouse triode.
1) Plate; 2) grid; 3) cathode;
4) heater.

upper section of the plate rod, which comes out above the disc seal, a cylinder is connected; it forms the inner cylinder of a coaxial line. The grid is made of woven molybdenum mesh, soldered to a ring which is in turn soldered to the grid contact disc. The indirectly heated tube cathode takes the form of a cylinder closed at one end, with the closed end oxide-coated. The cathode cylinder is welded to a disc which is insulated from the metal shell of the tube by a mica disc. There is a considerable capacitance between the cathode disc and the shell (on the order of 100 $\mu\mu$f), and as a result, the shell serves as the RF contact of the cathode. In order to apply a direct voltage to the cathode, it is necessary to use leads that pass through a glass stem and that are connected with pins in the base. The heater is also let out through the stem; a double-helical-type heater is used.

The small separation between grid and cathode (on the order of 0.1 mm) makes it necessary to utilize a rather fine grid, which guards,

first, against the appearance of an "island" effect and, second, provides a small transfer capacitance. Thus, lighthouse triodes have, as a rule, a rather high amplification factor, on the order of 40-60.

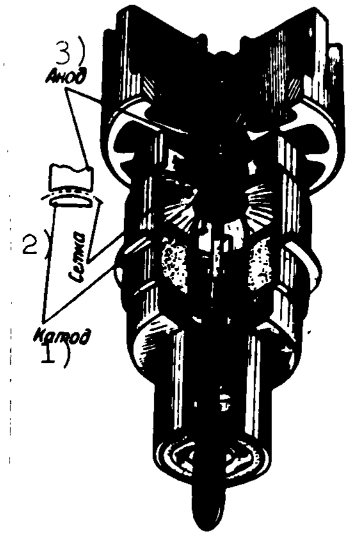

Fig. 12-21. Metal-ceramic transmitting triode.
1) Cathode; 2) grid; 3) plate.

Lighthouse triodes are designed to operate in the decimetric, and upper end of the centimetric (to 8 cm) bands.

Considerably more power may be delivered by metal-ceramic microwave transmitting triodes (Fig. 12-21). The envelope of such tubes consists of metal cylinders and a vacuum ceramic. The tube electrodes are spherical in shape; the cathode has a convex working surface, while the plate has a concave working surface. In contrast to the plane electrodes utilized in lighthouse tubes, these electrode shapes have high stability under thermal deformation, which makes it possible to increase the operating temperatures of the electrodes in service.

The cathode and grid contacts are formed by external metal cylinders that are in contact with the cylinders of the external coaxial line; they are separated from each other by a ceramic element. The plate rod is sealed into the upper ceramic disc and has a screw by

61-102

which the air-cooling radiator is mounted. The indirectly-heated oxide-coated cathode is mounted with the aid of clamps to a crimped nickel

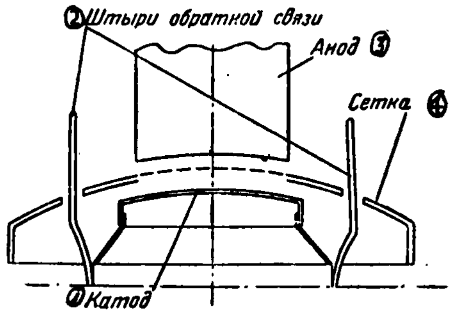

Fig. 12-22. Diagram of pin feedback for a metal-ceramic triode designed to operate as a self-excited oscillator.
1) Cathode; 2) feedback pins; 3) plate; 4) grid.

cylinder the lower end of which is welded to the cathode-lead cylinder. The upper internal section of the cathode cylinder is joined to a ceramic disc, inside of which is sealed a metal tip which at the same time forms a lead of one of the heater rings. The second heater ring is connected to the cathode. This cathode-unit design permits a precise establishment of the grid-cathode spacing in the evacuated tube by

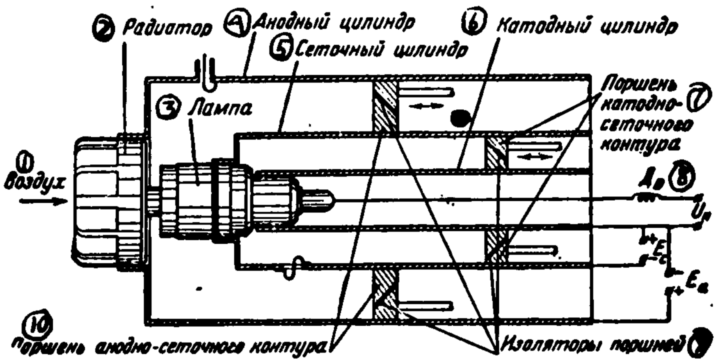

Fig. 12-23. Oscillator using metal-ceramic tube.
1) Air; 2) radiator; 3) tube; 4) plate cylinder; 5) grid cylinder; 6) cathode cylinder; 7) cathode-grid-loop plunger; 8) choke; 9) plunger insulator; 10) plate-grid-loop plunger.

placing the crimped cathode cylinder under tension or compression. This adjustment is made on the basis of a measurement of the grid-cathode capacitance.

A grid in the form of a wire mesh is soldered to a nickel spherical ring that is welded to an outside cylinder that forms the grid contact.

Owing to the possibility of forced-air cooling of the plate, and the high heat-resistance of the metal-ceramic envelope and interior tube elements, tubes of this design may deliver considerable RF power (up to 100 watts or more) in the decimetric range. The lower wavelength limit is about 6 cm.

As in the case of lighthouse tubes, in metal-ceramic tubes the surface conductance of the high-frequency leads is of great importance. In order to increase this conductance, the metal surfaces of the tube are silver-plated.

Proper operation of cathodes is of great importance in microwave-tube service. For large electron transit times (relative to the period of the alternating voltage on the grid) in the grid-cathode space, some of the electrons leaving the cathode during the positive half-cycle of voltage on the grid do not succeed in passing the grid and are returned to the cathode during the negative half cycle. As a result, electronic bombardment of the cathode occurs, leading to overheating. When a tube is operated under steady-state conditions, a considerable amount of power is developed at the cathode; its magnitude depends upon the mean value of cathode current and upon the working frequency. This power turns out to be so large that it is necessary to drop the cathode heating voltage in order to maintain stable temperature conditions at the cathode. In some cases, it turns out to be possible to disconnect the heater voltage altogether after the tube has established

its operating regime. The type of tubes described are used in microwave oscillator circuits as both independently excited and self-excited oscillators. With self-excited oscillators, it is necessary to obtain feedback from the tube output circuit to the input circuit. This feedback may be obtained either with the aid of the coupling loop between circuits of the oscillator itself, or through the tube transfer capacitance. When triodes are operated in grounded-grid circuits, the plate-cathode capacitance, which is very low in metal-ceramic tubes, may serve as the transfer capacitance. Thus, in several metal-ceramic tubes specially designed to operate as self-excited oscillators, an artificial increase in transfer capacitance is effected by means of internal coupling pins connected to the cathode and passing through special holes in the grid ring into the space near the plate (Fig. 12-22).

An oscillator circuit using a metal-ceramic tube is shown in Fig. 12-23. The RF leads of the plate, grid and cathode are connected to three coaxial cylinders. The inside surface of the plate cylinder and the outside surface of the grid cylinder, shorted by a plunger, form the output plate-grid circuit, while the inside surface of the grid cylinder and the outside surface of the cathode cylinder, also short-circuited by a sliding plunger, form the input grid-cathode circuit. A coupling loop is used for feedback between the output and input circuits; it passes through a window in the grid cylinder. The oscillator is tuned by moving the plungers, which are fastened to special rods. The plungers consist of two parts separated by thin insulators, which are necessary to isolate the DC power-supply circuits. Since the capacitance between the insulated sections of the plungers is quite large, it presents no substantial resistance to microwave currents. In such cases, the isolating insulators are installed not in the plungers, but where the cylinders are connected to the tube

contacts (capacitive contact). In the case of the heater supply, RF chokes are inserted; they guard against a loss of RF power owing to

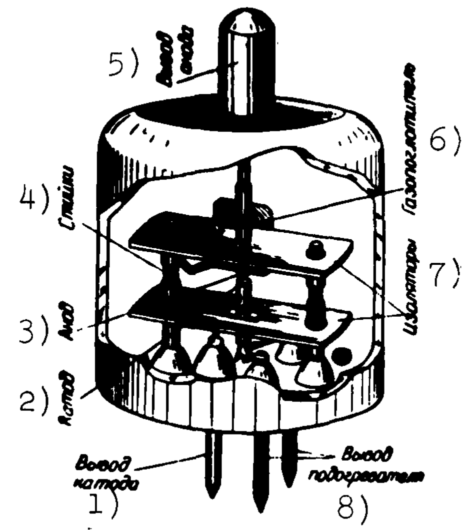

Fig. 12-24. Diode with pin seals.
1) Cathode lead; 2) cathode;
3) plate; 4) support; 5) plate
lead; 6) getter; 7) insulators;
8) heater leads.

microwave currents induced in the heater. Oscillator circuits using lighthouse tubes are designed in similar fashion.

Diodes are used for detecting modulated microwave signals; these diodes must meet the same structural requirements as do microwave triodes. One possible design for such a diode is of the lighthouse-tube type, while another takes the form of a glass diode with a button stem and a pin plate lead led out through the top of the envelope (Fig. 12-24). In both tube designs, where the working electrode surfaces are plane, the plate-cathode capacitance depends, in practice, solely upon the magnitude of the working surfaces of the electrodes, and their separation, and have values on the order of 0.2-0.3$\mu\mu$f. The small cathode-plate spacing provides a comparatively short electron transit time. For this reason, such diodes may be used to detect signals at wavelengths to 10 cm.

61-10

Chapter Thirteen

KLYSTRONS, MAGNETRONS, AND TRAVELING-WAVE AND BACKWARD-WAVE TUBES

13-1. THE PRINCIPLE OF ELECTRODYNAMIC CONTROL OF AN ELECTRON BEAM. THE KLYSTRON.

In 1932, Professor D. A. Rozhanskiy proposed, in principle, a new method for the excitation of high-frequency oscillations, which made possible the subsequent design of the electron tubes that have come to be called klystrons. The distinctive feature of this method is the acceleration of electrons by a constant electric field, accompanied by the action of a changing high-frequency field, which divides the uniform electron stream into separate groups (bunches) of electrons.

Let us examine the principle by which the uniform stream of electrons is changed into a beam of varying density.

Electrons escaping from the hot cathode k and focused into a narrow beam are accelerated by a high positive potential on anode a_1, and pass through two grids c_1 and c_2 located on their paths, at high velocities v_0 (Fig. 13-1). In the absence of a potential difference be-

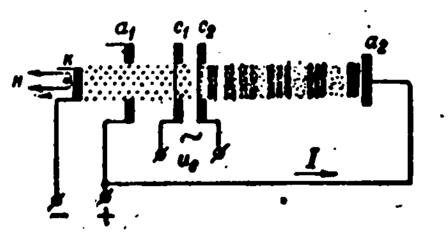

Fig. 13-1. Changing field of grids groups a uniform stream of electrons.

tween the grids, the electrons without losing velocity, strike the second anode a_2. In this case, the electrons move from the cathode to anode a_2 in a uniform stream, and direct current flows in the external circuit. If we now apply a changing RF voltage to the grids and, thus, set up a longitudinal high-frequency field in the space between the

grids, during the positive half-cycle, the electrons will be accelerated by the electrical field of the grids, while during the negative half-cycle, they will be delayed. As a result, the electrons, approaching the grids at uniform velocities, leave the grid region with velocities varying from $v_0 - \Delta v$ to $v_0 + \Delta v$. When the potential difference between the grids passes through zero values, the velocity v_0 of the electrons will not be changed. After leaving the grid region, the electrons that have a velocity v_0 approach the electrons that have previously passed through the grids at a velocity below v_0 (to $v_0 - \Delta v$). In turn, accelerated electrons having a velocity higher than v_0 (to $v_0 + \Delta v$), overtake the electrons that are moving at a velocity v_0. At a certain distance from the grids, the electrons that pass through the grid region with a time corresponding to the minimum-to-maximum region of a cycle will form a "bunch" of electrons, with a high density. Electrons that have passed through the grid region over a portion of the cycle corresponding to a change in voltage from maximum to minimum leave the grids with gradually diminishing velocities and scatter, resulting in a decrease in the density of the electron beam. If the plate a_2 is located at the point where the densest groups of electrons form, these groups, striking the anode, will set up a pulsating current in the anode circuit. Since the groups of electrons arrive at the anode a_2 with a frequency equal to the frequency of the alternating voltage on the grids, the current in the plate circuit should pulsate with the same frequency. If we introduce a tuned circuit (Fig. 13-2) into the circuit of the anode a_2, and tune it to the frequency of the current pulsation, there will occur undamped oscillations at the same frequency. Such a device will operate as either an amplifier or an oscillator of the self-excited type. If a portion of the RF energy of the tuned circuit is fed back to the grids with appropriate phase, it is possible, using

this method, to set up a self-excited oscillator circuit.

The method of converting a stream of electrons at constant density into a stream with a continuously varying density is widely utilized in the electronic devices that are called klystrons.

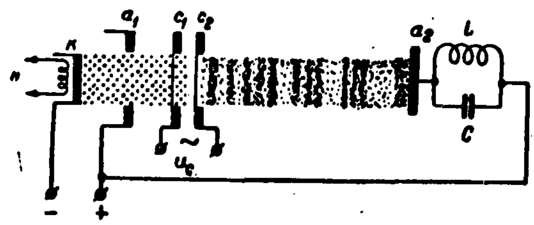
Fig. 13-2. Basic circuit of klystron.

Klystrons are chiefly utilized for service in the centimetric and millimetric bands. Thus, the tuned circuits used with them are cavity resonators, most frequently of toroidal shape, and located at the point where the electrons gather into the densest groups, i.e., at the location of anode a_2. As they pass between the cavity grids, the bunches of electrons induce alternating electrical charges on them; as they flow along the interior walls of the resonator from the first grid to the second and back, they form an alternating electric current of high frequency. As a result, an RF electromagnetic field forms in the resonator cavity; the energy of this field is coupled out to the point at which it is to be used by means of a coupling loop. Losses of RF energy in the resonator are compensated through the transmission of kinetic energy by the bunches of electrons to the electric field of the resonator through the charges induced in the resonator. Electrons that have given up a considerable portion of their kinetic energy for the excitation and maintenance of oscillations in the resonator leave the resonator at low velocities, and fall onto a collector which is at a positive potential with respect to the cathode. In the last analysis,

the energy obtained by the electrons from the constant electric field set up by the voltage source is given up by the electrons to the high-frequency field of the output tank circuit; in other words, what occurs

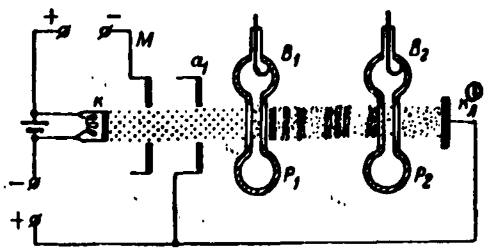

Fig. 13-3. Basic circuit of double-cavity klystron. 1) k_1.

is a conversion of DC energy into the energy of high-frequency electrical oscillations.

Thus, the basic distinction between klystrons and tubes with electrostatic control lies in the fact that the beam of electrons is controlled not by regulating the velocity of the electrons in the cathode-grid space as in, for example, a triode, but by: a) acceleration of the electrons by a constant electrical field in the space between the cathode and anode a_1; b) changing the velocity of the electrons in the space between the grids; c) grouping the electrons in the space between grid c_2 and the resonator (the bunching space). The advantage of this method of control lies in the fact that direct control of the electron beam is performed following acceleration of the electrons by the constant electric field in a space where no high-frequency field is acting. Thus, the electrons traverse the distance between the grids in a klystron with exceptionally high velocities and, consequently, in a very short time, which permits a considerable increase in the upper working-frequency limit.

In order to decrease the electron transit time, and to obtain as high an oscillation frequency as possible, the spacing between the grids

of the resonator is made extremely small, and the electrons obtain high velocities (on the order of $5 \cdot 10^9$ cm/sec) before entering into the space between the grids. The very small spacing between the grids, however, increases circuit capacitance.

This principle is the basis for the so-called two-cavity (transit) klystron that uses two resonators (Fig. 13-3): an input resonator p_1 and an output resonator p_2. The electrons of a bundle are accelerated by anode a_1 and strike the opening in the input resonator p_1; a changing high-frequency potential difference is applied between the grids of this resonator from an external oscillator by means of a coupling loop V_1 introduced into the resonator cavity. The high-frequency field, acting between the grids of the input resonator, converts the uniform electron beam into a beam of electrons of varying density; almost no energy is consumed in this conversion process, since in a single cycle the number of retarded electrons that yield kinetic energy to the field equals the number of accelerated electrons that gain just as much energy from the field. Leaving resonator p_1, the electrons move toward the output resonator p_2. As they approach resonator p_2, dense groups of electrons are formed, which then pass through the grids of this resonator at a rate of 1 per cycle, i.e., at the frequency with which the field between the resonator p_1 grids changes.

Moving in the space between the grids of resonator p_2, the groups of electrons expend a considerable portion of their kinetic energy in exciting oscillations in the resonator; if the electron groups are short in length and traverse the space between the grids when the retarding field is at a maximum, i.e., at the peak value of the alternating voltage between the grids, they yield the maximum amount of energy to the resonator; if the groups are longer, and do not pass through the space between the grids when the field is at a maximum, many electrons

yield a small amount of energy to the resonator, since they move in a retarding field at lower field strength. When they pass through the space between the grids during an accelerating half cycle, groups of electrons carry energy away from the resonator field; as a result, such groups of electrons facilitate damping of the oscillations in the resonator. Thus, the transit time for groups of electrons in the space between grids of resonator p_2 is so chosen that the majority of electrons pass through the grids at retarding-field values close to the peak value.

The kinetic energy taken up by resonator p_2 from the groups of electrons is converted into oscillatory energy of the electromagnetic field within the resonator, and then, by way of the coupling loop introduced into the resonator cavity, is applied to the next stage.

The electrons that have given up a portion of their kinetic energy to the output resonator and have thus suffered a decrease in velocity strike the collector k_1.

Since there is no external field in the space between the input and output resonators, and in many klystrons the outer surfaces of their resonators lie at different constant potentials, for convenience, both resonators are interconnected or combined into a single structure, while maintaining the distance between the output grid of the input resonator and the input grid of the output resonator that is required for electron bunching. In certain designs, the anode a_1 serves as the input resonator to which a high positive potential is applied.

In the two-cavity klystron, oscillations may be excited by accurately tuning the input and output resonators to the given frequency, which is very difficult to do, since both resonators have sharp resonance curves. Both resonators are tuned to the given frequency, as a rule, by changing their volumes or the separations between grids with the

aid of special devices.

Two-cavity klystrons are utilized advantageously as voltage amplifiers or RF oscillators in the centimetric range. The power gains of such klystrons do not exceed 20. For increased gain, several klystrons

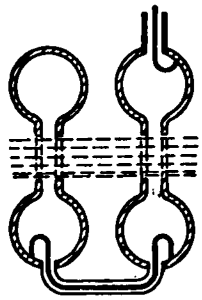

Fig. 13-4. Feedback with the aid of a coaxial line.

are connected in series, or multicavity klystrons are used.

In a klystron oscillator, weak feedback is introduced between the resonators with the aid of coupling pins or loops, connected by coaxial lines (Fig. 13-4), or directly by means of a special aperture in the wall that separates the input resonator from the output resonator. With very heavy feedback, the amplitude of the voltage between the grids of the input resonator is so great that strongly accelerated and strongly retarded groups of electrons are formed, resulting in "electron clusters" that form a very short time after leaving the input resonator, in direct proximity to it. Then, as they approach the output resonator, these clusters become ever more dispersed. As a result, the efficiency of the klystron drops, or the klystron totally ceases to oscillate.

The maximum klystron efficiency attainable in practice is about 20 per cent, although theoretically, it may be about 58 per cent. This discrepancy is explained by the fact that from 50 to 75 per cent of the electrons of a bundle are scattered, and, falling on the walls of the klystron, give up their kinetic energy as heat and thus perform no use-

ful work. In addition, electron clusters tend to disperse owing to their mutual repulsion, which also decreases efficiency.

Oxide-coated cathodes are utilized in klystrons, and their emission current (depending upon the type of klystron) amounts to 80 to 500 ma or more. The accelerating voltages acting between the cathode and input resonator are different for different types of klystrons: in low-power klystrons it varies from 1000 to 1200 v, in medium-power klystrons, it is on the order of 3 kv, and in high-power pulsed klystrons, it is from 10 to 13 kv. Here, the useful power in continuous operation reaches hundreds of watts, and in pulsed operation, it ranges from several tens of kilowatts to tens (and somewhat more) of megawatts.

A drawback to the two-cavity klystron as an amplifier is the high internal noise level resuting from the effect of fluctuations in the electron beam. This causes "noise" voltages at the input resonator which are superimposed upon the signal voltage, and are amplified together with the signal by the output resonator.

In the two-cavity klystron oscillator, oscillations may be excited only where there are tuned circuits and where the supply voltages are properly chosen. In practice, however, it is quite difficult to tune circuits to resonance and to choose the necessary supply voltages. In addition, detuning of a klystron oscillator frequently occurs in operation owing to changes in supply voltages, which causes changes in the velocity, in the repetition rate for electron bunches, and in the point of bunching, and these effects can also occur as a result of changes in the geometric dimensions of the resonators caused by their heating. This occurs especially vigorously in high-power klystrons, in which the dimensions of the resonators and the separation between their grids change noticeably owing to heating during operation, leading to corresponding changes in frequency of oscillations. As a result, the oscilla-

tions excited in a klystron become unstable.

Some of these drawbacks to the two-cavity klystron are removed in the considerably more common reflex klystron, in which there is a single resonator in place of the two resonators, which eliminates the need for tuning of tank circuits, since the function of the input and output resonators is combined into a single resonator, and tuning is accomplished solely by the choice of supply voltages.

13-2. THE REFLEX KLYSTRON.

The reflex or single-cavity klystron, suggested in 1940 by the Soviet engineer V.F. Kovalenko, consists of a hot cathode K, a control

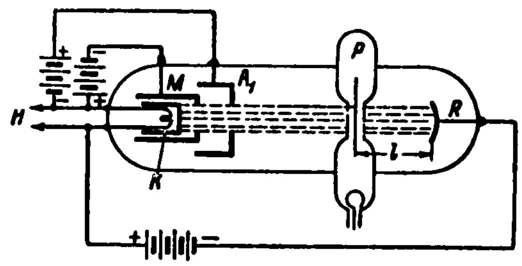

Fig. 13-5. Schematic diagram of a reflex klystron.

electrode M, an anode A_1, a cavity resonator P, and a repeller R (Fig. 13-5). The electrons leaving the cathode are focused with the aid of the control electrode M and focused into a quite-narrow beam and accelerated by the action of the high positive potential on anode A_1. The control electrode M, to which a small positive or negative potential is applied, also serves to regulate the magnitude of the current in the electron beam through changes in potential. The shape of the cathode is so chosen that when the electrons produced by it are focused, the cross-section of the beam of electrons obtained is circular.

The beam electrons, accelerated by the high anode potential, enter the space between the resonator grids at high initial velocities v_0, are

separated by the resonator field into separate bunches which move at different velocities toward the electrode R, which is negative with respect to the cathode, and located a distance $\underline{l}$ from the resonator. The bunches of electrons leaving the resonator enter a strong retarding field and, without reaching the electrode R totally lose their velocity. In this case, electron bunches possessing higher velocities penetrate deeper into the retarding field and, consequently, come closer to electrode R than do electron bunches having lower velocities. After losing their velocities, the electron bunches begin to travel in the opposite direction, i.e., back toward the resonator, since in this direction, the electric field is accelerating for them. Thus, as it were, the electron bunches are reflected from the electrode R and are not incident upon it. Thus, electrode R is frequently called the repeller. In order to increase the focusing action of the repeller, it is made concave with respect to the resonator.

Bunching of electrons takes place along the path from the resonator to the repeller and back toward the resonator (as in the two-cavity klystron). At a given distance $\underline{l}$ between the resonator and the repeller, it is possible to select an accelerating voltage for anode A_1 and a retarding voltage for repeller R such that electron bunching will occur at the input to the resonator at the instant of time that corresponds to the maximum retarding field between the resonator grids. At these instants, bunches of electrons will yield a considerable portion of their kinetic energy to the high-frequency field of the resonator.

Spent electrons are intercepted by the first resonator grid from the cathode or by anode A_1.

Thus, in the reflex klystron, the same resonator functions as the input resonator p_1 and the output resonator p_2 of the two-cavity klystron and, consequently, there is no need for tuning circuits into

resonance.

In contrast to the two-cavity klystron, in which electron bunching takes place in the space between the input and output resonators in the absence of an electric field, in the reflex klystron, electron bunching takes place in the space between the resonator and the repeller, in which a strong electric field acts.

In practice, it proves convenient and advantageous to excite one type of oscillation or another in a klystron by changing the repeller voltage U_R, without changing the accelerating voltage U_A. This method of controlling a klystron, also called electronic tuning, is extremely responsive, and involves no power losses, since the repeller consumes almost no power.

Some versions of reflex and two-cavity klystrons do not use the anode A_1. In such designs, the accelerating voltage U_A is applied directly to the resonator P.

In the reflex klystron, oscillations are not excited for any values of beam current, but only for those values that exceed a certain minimum, frequently called the starting current. If the current in the beam is lower than the starting current, oscillations will not be excited in the klystron, since the beam energy is insufficient to excite and maintain oscillations.

Reflex klystrons are widely used in low-power oscillator tubes for the centimetric band in superheterodyne receivers, for radio relay lines, and under laboratory conditions, for developing and adjusting centimetric-band equipment.

Construction and tuning of reflex klystrons are simple. The frequency of oscillations generated is maintained constant, as a rule, by applying to the anode (or resonator) and repeller heavily regulated voltages. For this purpose, a common source is used to supply voltage

to the circuits of the anode A_1 and repeller R.

13-3. MOTION OF ELECTRONS IN INTERSECTING ELECTRICAL AND MAGNETIC FIELDS.

The principle of controlling a beam of electrons with the aid of electrical and magnetic fields is utilized in electronic devices that are called magnetrons. The simplest magnetron is a two-electrode tube with a filament-type cathode located on the axis of a cylindrical anode. The essential difference distinguishing the very simple magnetron from a common diode is the fact that the interelectrode space of the magnetron lies in a constant magnetic field whose lines of force are directed parallel to the cathode and anode and, consequently, perpendicular to the electric field-strength vector of the anode. For this reason, electrons given off by the hot cathode simultaneously experience effects from the electric and magnetic fields. Under the action of the electric field of the anode, the electrons attempt to move toward the anode in radial directions perpendicular to the magnetic field, while under the action of the magnetic field, their rectilinear motion is turned into curvilinear motion.

Let us consider the path of motion of electrons in the space between the cathode and anode of a very simple magnetron in the presence of different values of magnetic field strength H and constant plate voltage U_a.

Let electrons leave the magnetron cathode at zero initial velocity ($v_0 = 0$). If in the space between the cathode and anode, the electric field is not distorted by the space charge, and there is no magnetic field, the electrons will move, under the action of the accelerating electric field of the anode, along radial paths, and they will all be incident upon the anode. As a result, there will be a maximum anode current I_a in the external circuit of a magnetron, whose

value is limited by the emission power of the cathode (Fig. 13-6a). If in the space between the cathode and the anode there is a relatively weak magnetic field directed parallel to the cathode, the electron paths will curve, while the anode current will retain its initial value, since in this case all the electrons will be incident upon the anode (Fig. 13-6b). An increase in the magnetic field strength will produce still greater curvature of the electron trajectories, but the value of the anode current will remain unchanged. The curvature of the paths will be the greater the lower the anode voltage, and the higher the magnetic field strength. At a certain value of field strength which is called the critical value (H_{kr}), the paths will curve so much that the electrons will only touch the anode and then will return to the cathode (Fig. 13-6c). In this case, the anode current will at once drop to zero, since now not a single electron will be incident upon the anode. If the magnetic field strength is increased above the critical value, the electrons that do not reach the anode will be turned back by the strong magnetic field to the cathode (Fig. 13-6d) and the anode current will continue to equal zero.

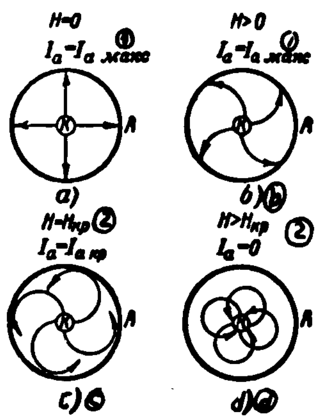

Fig. 13-6. Change in paths of motion of electrons in magnetron. 1) Max; 2) H_{kr}.

Thus, where $H < H_{kr}$, the plate current has its maximum possible

value, while at H $\gg$ H_{kr}, the plate current at once ceases, i.e., the slight change in magnetic field strength near the critical value causes a sharp change in the anode current. As a result, considerable current

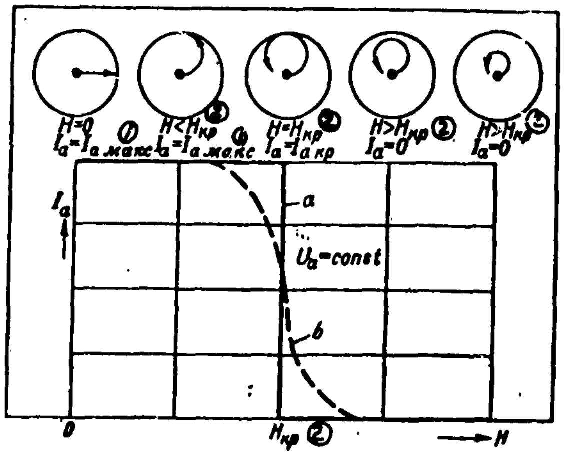

Fig. 13-7. Magnetron plate current as a function of magnetic field strength.
1) Max; 2) H_{kr}.

pulses appear in the anode circuit of a magnetron; in the presence of a tuned circuit, these pulses may excite powerful undamped oscillations in the tank circuit.

The dependence of a magnetron's plate current upon the magnetic field strength $I_a = f(H)$ may be represented graphically by plotting along the X axis the values of H, and along the Y axis, the anode-current values I_a (Fig. 13-7). In the ideal case, the curve $I_a = f(H)$ takes the form of a broken line consisting of a line parallel to the X axis from 0 to H_{kr}, a vertical line at H = H_{kr}, and a line coinciding with the X axis along the segment H $>$ H_{kr} (Fig. 13-7, line a). In reality, such a sharp change in plate current with a slight change in magnetic field strength is not observed (Fig. 13-7, curve b). The factors causing the upper and lower bends in the curve are the absence of strict parallelism between the magnetic field and the axis of the magnetron, distortion of the electric field at the ends of the electrodes, eccen-

tricity in the location of the cathode, etc.

Above the curve shown in Fig. 13-7, we have illustrated the changes in electron paths and plate current with changing magnetic field strength with the anode voltage constant, corresponding to curve a.

The relatively sharp variation in anode current caused by a slight change in the critical magnetic field strength is manifested by the possibility, in principle, of controlling the anode current by correspondingly slight changes in magnetic field strength and, consequently, the possibility of utilizing the magnetron as an oscillator and amplifier.

The method of controlling the anode current directly with the aid of the magnetic field has not been employed, however, since it requires extremely complicated equipment. It is more advantageous in practice to control the anode current in a magnetron by changing the anode voltage while maintaining the magnetic strength constant, i.e., to utilize the magnetron curve $I_a = f(U_a)$ at $H = \text{const}$, which is reminiscent of the shape of the curve for a common diode (Fig. 13-8).

The value of anode voltage at which anode current will appear in a magnetron for a given magnetic field strength is called the critical anode voltage for a magnetron $U_{a.kr}$. It follows from the curve that the greater the magnetic field strength, the higher the critical anode voltage should be.

The very simple magnetron described above is called a magnetron with a nonsegmented (solid) anode or a single-cavity magnetron. In such a magnetron, use is made of the oscillation of electrons moving in the space between the cathode and anode under the action of the electrical and magnetic fields; the electrons move in cycloids.

If a magnetic field with strength H_{kr} is set up in a magnetron

with cylindrical electrodes, and between the anode and cathode there is connected a tuned circuit, tuned to the frequency at which the electrons will describe cycloids in the magnetron, when anode voltage is applied

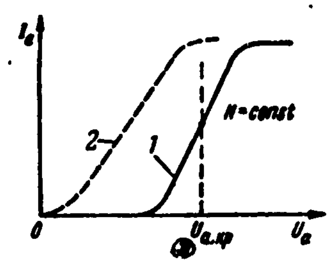

Fig. 13-8. Volt-ampere characteristic of magnetron.
1) $H > 0$; 2) $H = 0$;
3) $U_{a.kr}$.

to the magnetron, oscillation will automatically occur (the magnetron is self-excited from the voltage surge), where the period of oscillation will equal the transit time of the electrons along the trajectories described by them.

Single-cavity magnetrons are not in wide use, since it is impossible to obtain large RF power or good efficiency from them; efficiency normally does not exceed 10 per cent. The reason for this is that in the single-cavity magnetron, it is impossible to provide proper phasing of the electrons and, in addition, it is not possible to clear the working space completely of spent electrons which, after transmitting their energy to the tuned circuit, still remain, and again begin to be accelerated by the oscillating field, which expends energy upon them uselessly. If the magnetron anode is split into two semicylinders and the tuned circuit is connected directly between them (Fig. 13-9), more powerful oscillations will appear in such a magnetron. Such magnetrons are called split-anode magnetrons. Two basic types of oscillations are excited in them: the previously considered electronic oscillations at cyclotron frequency, very short in wavelengths, and low in power; and

the more powerful, longer wavelengths which are basically determined by the tank-circuit parameters, and are independent of the cyclotron frequency of the electronic oscillations. Oscillations of the second

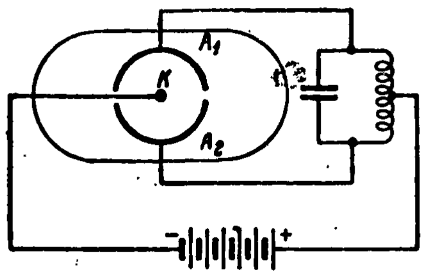

Fig. 13-9. Basic circuit for split-anode magnetron.

type are caused by the interaction of electrons with the high-frequency alternating field acting tangentially in the slots between the edges of the anode semicylinders.

13-4. MULTICAVITY MAGNETRONS.

Split-anode magnetrons are not in common use either, since they are uneconomical, and do not deliver sufficiently powerful and stable oscillations at superhigh frequencies. At the present time, for decimetric and centimetric-band oscillators, wide use is made of so-called multicavity magnetrons, that are more stable in operation, more economical, and that deliver greater amounts of RF power at superhigh frequencies either in pulsed operation or in continuous operation.

In multicavity magnetrons, anode units are used that have an even number (from 8-40) of resonators arranged around the circumference, equal distances apart. Figure 13-10 shows the arrangement of a multicavity magnetron: in the central cavity of the anode unit, there is a cylindrical cathode, while covers are soldered to the ends of the unit. The magnetic field of a magnetron is set up by a permanent magnet (or electromagnet) having pole pieces whose diameter is somewhat larger

than the diameter of the central cavity of the anode unit. In certain magnetron designs, the pole pieces simultaneously form the covers for the plate unit, which to a still larger degree concentrates the magnetic

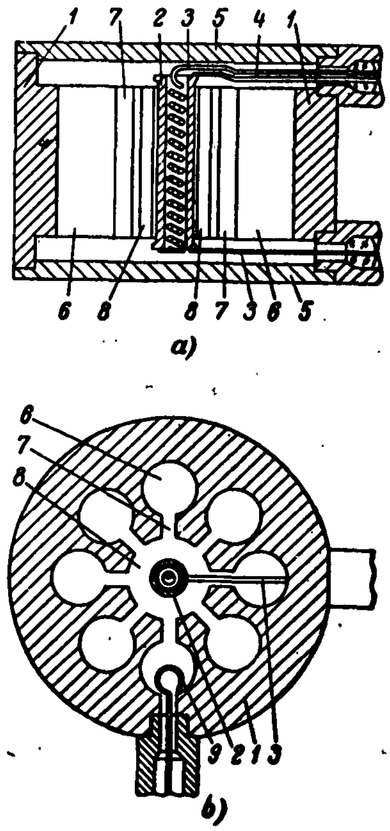

Fig. 13-10. Multicavity magnetron. a) Longitudinal section; b) lateral section; 1) anode unit; 2) cathode; 3) heater leads; 4) cathode lead; 5) covers; 6) resonator cavities; 7) resonator slots; 8) interaction space; 9) coupling loop.

field in the space between the cathode and the anode unit (the interaction space), and makes it possible to employ small permanent magnets.

For better cooling of the magnetron, the outside side surface of the anode unit has fins past which air is blown.

A high anode voltage is applied directly to the housing of the anode unit, which is grounded. The magnitude of this voltage depends upon the type of magnetron and varies from 1 kv to several tens of

kilovolts.

In a multicavity magnetron, as in a split-anode magnetron, oscillations occur owing to the interaction of electrons leaving the cathode

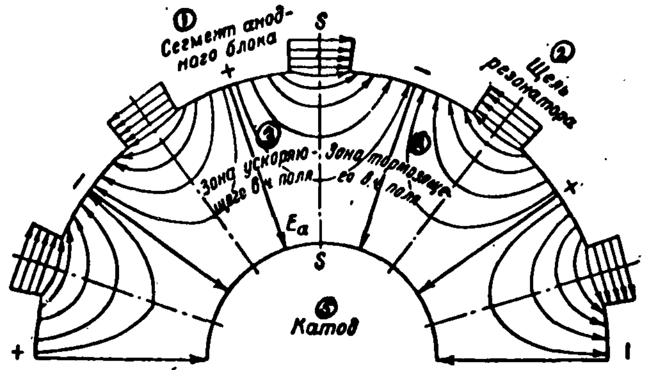

Fig. 13-11. Instantaneous distribution of high-frequency field of resonator and constant field of anode in multicavity magnetron.
1) Segment of anode unit; 2) resonator slot; 3) zone of accelerating high-frequency field; 4) zone of retarding high-frequency field; 5) cathode.

with the high-frequency electrical field of the resonators, each of which is a tuned circuit with a capacitance C_r, concentrated mainly at the edges of the resonator, and an inductance L_r distributed along its inner surface. The nature of the distribution of the lines of force of the high-frequency electric field of a magnetron is illustrated in Fig. 13-11 for one instant of time. The RF currents flowing along the inner surfaces of the resonators form a high-frequency magnetic field whose lines of force pass through the cavities of adjacent resonators. As a result, there is inductive coupling between all resonators of a magnetron (Fig. 13-12), which makes it possible to take RF power from a magnetron by introducing a coupling loop into a single resonator alone.

When a magnetron is in operation, an alternating high-frequency voltage acts between the segments of the anode unit; the high-frequency

voltages of neighboring segments are shifted in phase by 180°. With such oscillations, in any of the half cycles half of the segments, i.e., every other segment, are at a negative voltage. In the next half

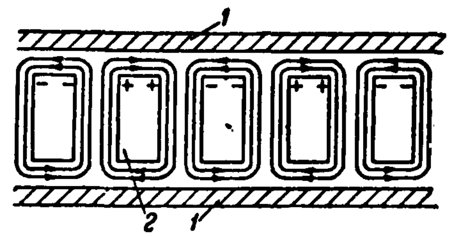

Fig. 13-12. Inductive coupling between resonators (view of magnetic fields from cathode). 1) Anode-unit cover; 2) anode segments.

cycles, the polarity of the segments changes, again alternating. Thus, the RF field in neighboring resonators is always opposite in direction, i.e., shifted by π [rad] (180°). Thus, such oscillations have come to be called π-mode oscillations. In multicavity magnetrons, the π-mode is the basic type of oscillation.

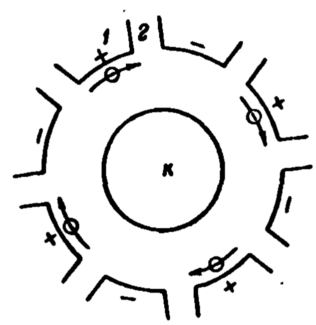

Fig. 13-13. Operating principle of multicavity magnetron. 1) anode-unit segments; 2) resonator slots; K) cathode.

Electrons leaving the cathode possess the required velocities owing to the difference of potentials between the cathode and the plate unit. The magnetic field directed perpendicular to the electric field

and parallel to the cathode causes the trajectories of the electrons to curve so that electrons move along the working surface of the anode unit, between whose segments the RF voltage acts.

For the sake of simplicity, let us suppose that there is a single electron opposite the center of each segment of the anode unit that is at a positive potential (Fig. 13-13). If these electrons are forced to move with the same initial velocity along the working surface of the anode unit (for example, in a clockwise direction), during the first half cycle of each cycle as they pass by the resonator slots, they will move in a retarding RF field from the resonators, while in the second half cycle, they will move in an accelerating field. When moving in a retarding field during the first half cycles, the electrons give up a portion of their kinetic energy to the RF field of the resonators, while in moving in the accelerating field (during the second half cycles) they receive back from the RF field the same amount of energy. Thus, over entire cycles, the RF field of the resonators obtain no energy from the electrons, and the RF oscillations die out. The rate of motion of the electrons may be so chosen, however, that the transit time of the electrons between the centers of successive segments will equal a half cycle of the high-frequency voltage. In this case, the electrons will always move in a retarding high-frequency field and, interacting with it, will yield to it their kinetic energy acquired as a result of the action of the constant acceleration voltage between the anode and cathode. As a result, the oscillations in the magnetron remain undamped.

If at a given instant of time there is a single electron opposite the center of each segment of the anode unit that is under a negative potential, then a velocity may be chosen that will force the electrons to always move in an accelerating high-frequency field. In this case, the electrons will only take energy from the high-frequency field of

the resonators, and oscillation will cease.

Consequently, in a working magnetron, oscillation may be maintained provided that the electrons in their motion past resonator slots always encounter a retarding RF field, which is possible only for a particular electron velocity.

The principle of selecting electron velocities at which the electron transit time from the center of one segment to the center of the next segment equals a half cycle of the high-frequency-voltage oscillation is called the principle of synchronization.

Let us now suppose that a single electron is opposite the center of each segment. In this case, even where there is synchronization, in any half cycles, half of the electrons will move in an accelerating RF resonator field, opposite every other resonator, while the other half of the electrons move in a retarding high-frequency field also opposite alternate resonators. As a result, the energy obtained from the electrons by the RF field for any full cycles will equal the energy given to the electrons by the field. Thus, the magnetron will not oscillate. Consequently, in normal operation of a magnetron, the electrons should lie in those zones of the interaction space in which the high-frequency field is retarding, and should not be in zones where the high-frequency field is accelerating, i.e., the electrons, leaving the cathode and entering the interaction space, should be bunched only in zones with a retarding high-frequency field.

Let us consider the mechanism of electron bunching.

At any point of the interaction space, the voltage of the high-frequency field may be resolved into two components: a radial component and a tangential component. In zones under positive segments of the anode unit, the radial component is directed toward the cathode, while in zones under negative segments, it is directed toward the segments

(Fig. 13-14a). These zones are bounded by planes passing through the axis of the cathode and the centers of the slots. As follows from Figs. 13-14a and 13-11, the boundary planes pass alternately through

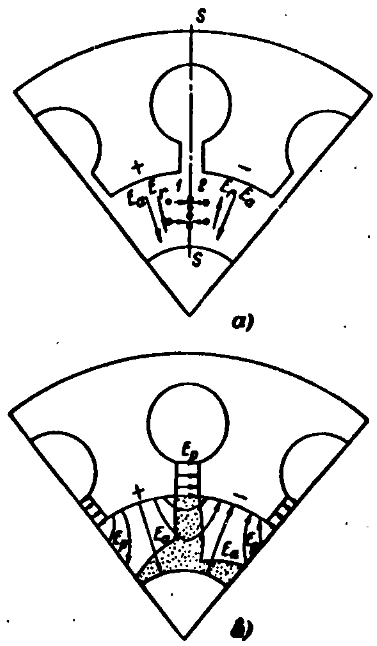

Fig. 13-14. Space-charge focusing by high-frequency fields.

the retarding and accelerating tangential RF fields.

Let us consider the motion of electrons to the left and to the right of one of these planes ss, which passes through a retarding tangential high-frequency field.

The electrons in the plane ss (Fig. 13-14a) are accelerated by a constant anode field and shifted by the magnetic field from left to right, while electrons to the left of plane ss are accelerated both by the constant anode field and by the radial component of the high-frequency field of the resonator. These electrons are also shifted from left to right by the magnetic field, i.e., toward plane ss, but extremely rapidly, since they have higher velocities. Thus, electrons to the left of plane ss overtake the electrons lying in this plane.

Electrons lying to the right of plane ss are accelerated by the

constant anode voltage, but are retarded by the radial component of the RF field, directed from the cathode to the segment. These electrons are also shifted from right to left by the magnetic field, but more slowly than electrons lying in plane ss, or those further to the left of this plane. Thus, electrons lying in plane ss and to the left of it overtake electrons lying to the right of this plane. As a result, in the zones of retarding RF field, the electrons come together (bunch) and form a narrow channel (blade), along which electrons are directed from the cathode to the anode (Fig. 13-14b). As follows from Fig. 13-14, the electron blades are formed in the zones of a retarding tangential field, in which the kinetic energy of the electrons is converted into energy of the high-frequency field, i.e., into energy of high-frequency oscillations. In zones with an accelerating high-frequency field, the blades do not form. Thus, the uniform electron cloud formed by the cathode in the interaction space is converted (focused) by the radial component of the RF field into an electron cloud with blades (into an "electron rotor") (Fig. 13-15). Since the retarding tangential RF field alternates, corresponding to every other resonator, the blades will also alternate, skipping every other resonator. Thus, if the number of resonators in the anode unit is N, the number of blades will be N/2. Thus, for example, for an 8-cavity magnetron, the number of blades is $N/2 = 8/2 = 4$.

The "electron rotor" rotates around the cathode with an angular velocity such that each blade moves from one slot of the anode unit to the next in a half cycle, i.e., in synchronism with the RF oscillations. Here the edges of the blades slip past the anode segments, while in each blade, the electrons move in cycloidal trajectories, going from cathode to anode, and yielding to the high-frequency fields of the resonators the energy obtained in the constant anode field. Depending upon the

direction of the magnetic field, the "electron rotor" may rotate clockwise or counter clockwise.

Let us now consider the nature of the motion of electrons in the

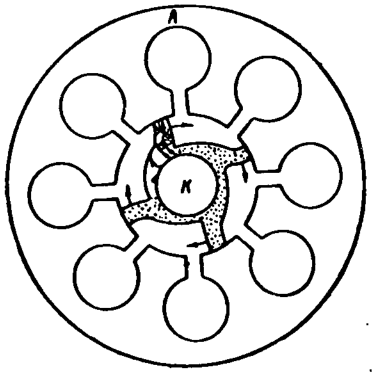

Fig. 13-15. Shape of space charge in magnetron oscillator.

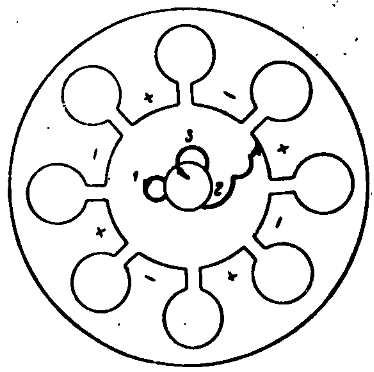

Fig. 13-16. Paths of electrons in magnetron oscillator.

electron blades.

Electrons leaving the cathode and accelerated by the constant anode field fall into a retarding tangential high-frequency field. The kinetic energy which they obtain in moving from cathode to anode owing to the accelerating anode field decreases, since a portion of

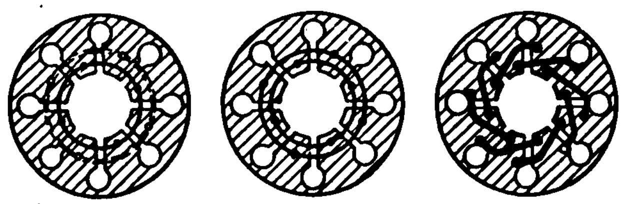

Fig. 13-17. Types of couplings between plate-unit resonators.

it is given over to the retarding high-frequency field. As the energy drops under the influence of the magnetic field, directed perpendicular to the electric field, the electrons begin to move toward the cathode, but do not reach it, since their energy becomes low. Remaining at a

certain distance from the cathode, the electrons again begin to move toward the anode, again delivering a new amount of energy to the high-frequency field. Under the influence of the magnetic field, the electrons again change their direction of motion, and remain still further from the cathode, since their energy becomes still less.

Thus, these electrons, obtaining energy from the constant anode field, and giving it up to the RF field, are constantly moving toward the anode, describing a set of cycloidal paths (trajectory 2, Fig. 13-16). Yielding a considerable amount of their energy to the RF field, the electrons, in moving further, strike the anode and are no longer involved in the process.

Let us now consider those electrons that upon leaving the cathode fall into an accelerating high-frequency field. Under the influence of the magnetic field, these electrons return to the cathode with energies obtained from the accelerating tangential RF field (trajectory 3, Fig. 13-16).

Such electrons do unfavorable work; they take energy away from the high-frequency field, and transfer it to the cathode, owing to which the cathode temperature rises.

Finally, in those zones of the interaction space where the high-frequency field strength is so small that it may be neglected, the electrons leaving the cathode receive no energy from the RF field, and under the influence of the magnetic field return to the cathode with zero energy (trajectory 1, Fig. 13-16). Such electrons perform neither useful nor harmful work.

Together with oscillations at the basic frequencies (π mode), oscillations may be excited in a multicavity magnetron at other frequencies as well that are close to the fundamental frequency (so-called parasitic oscillations). This also explains the arbitrary

transition of oscillations generated by the magnetron from the fundamental frequency to the spurious frequency, causing instability in the magnetron characteristic.

The effect of parasitic oscillations is eliminated by locating

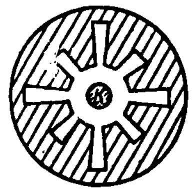

Fig. 13-18. Construction of anode unit with resonators of different dimensions.

the resonant frequencies further apart, which decreases the possibility of oscillations going over from the fundamental frequency to other frequencies; in this case, the energy of the constant electrical field is used solely to maintain oscillations in the basic mode. In practice, the resonant frequencies are separated by introducing special couplings between the sections of the plate unit (Fig. 13-17). These links are utilized in magnetrons designed to operate at frequencies of up to 10,000 Mc.

The presence of such links in a magnetron makes it possible not only to move the resonant frequencies apart and to establish a working regime with oscillations at the fundamental frequency without exciting oscillations at other frequencies, but also improves the stability of the characteristic and improves efficiency for the magnetron.

For SHF magnetrons, separation of oscillation modes is accomplished not with the aid of links, which involve considerable difficulties owing to the small dimensions of the resonators and links, but by using anode blocks having differing tuned circuits (diverse resonators) in which large and small cavities alternate; the cavities of different sizes will have different natural resonant frequencies (Fig. 13-18).

In this case, oscillations will occur with a frequency midway between the frequencies of the large and small cavities. Such rising-sun anodes give good mode separations and do not have the effects inherent in strapped anodes. Rising-sun magnetrons with a large number of resonators are very effective.

In certain designs for multicavity magnetrons, provision is made for special fittings for changing resonator inductance or capacitance by changing the volumes of the resonator cavities or slots by, for example, the utilization of special tuning rings or strips. In this way, it is possible to obtain more accurate tuning of a magnetron to the required working frequency. With this method, it is possible to obtain a change in frequency of about $\pm$ 7 per cent from the mean working value.

In high-power multicavity magnetrons operating under pulse conditions, indirectly heated oxide-coated cathodes are utilized, but their operating regimes in magnetrons differ from those in standard electron tubes. The distinction lies in the fact that the remaining kinetic energy of the electrons returned to the cathode, obtained by the electrodes from the RF accelerating field, goes to excite secondary-electron emission, and for additional heating of the cathode. This energy amounts, on the average, to 7-8 per cent of the energy applied to the anode circuit, and where the magnetron is operated at high frequencies (λ roughly 3 cm), it turns out to be sufficient to maintain the working temperature of the cathode with the heater supply turned off. In order to avoid overheating of the cathode in such magnetrons under working conditions, the cathode heating current is decreased or cut off altogether.

For firmer attachment of the oxide coating and weakening of the action of electron bombardment on the cathode, a nickel grid is used

that forms small cells into which the oxide mixture is pressed.

In magnetrons operating under continuous conditions, a tungsten cathode is utilized; it gives lower emission but is more stable in

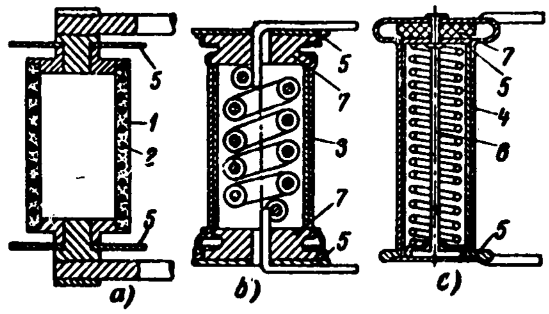

Fig. 13-19. Cathode designs for multicavity magnetrons.
a) Thorium-oxide; b) metallized wire gauze; c) oxide-coated metal gauze.
1) Thorium-oxide cylinder; 2) embedded tungsten helix heater; 3) nickel tube and grid; 4) nickel tube and grid with oxide mask; 5) screen; 6) support; 7) insulator.

operation. In magnetrons of this type, the useful power is small. The power developed at the magnetron plate, however, in operation under continuous conditions, is considerably greater than the mean power

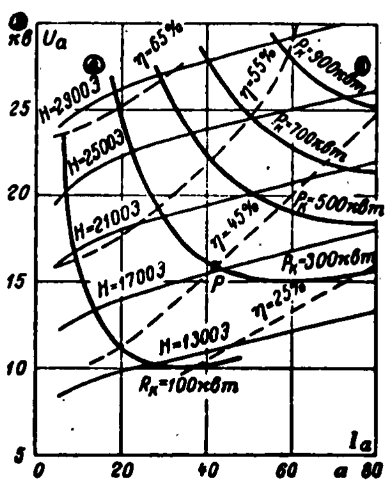

Fig. 13-20. Working characteristics of multicavity magnetron.
1) kw; 2) oersteds; 3) kv.

developed at the anode under pulsed operation. Thus, in such magnetrons, the plate is frequently water-cooled.

Figure 13-19 shows several types of cathodes utilized in multicavity magnetrons.

Multicavity magnetrons are very efficient. The reason for this is that the electrons do not penetrate deeply into the accelerating RF field and, consequently, they give up a small amount of energy to it. Conversely, electrons penetrate deeply into the retarding high-frequency field, and in moving from cathode to anode, give up to this field nearly all of the energies that they have gained from the constant electric field. As a result, in a multicavity magnetron, a considerable portion of the energy of the constant field is converted into RF-field oscillation energy, and efficiency increases.

The basic performance characteristics of a multicavity magnetron are the graphs showing the dependence of anode voltage upon anode current $U_a = f(I_a)$, corresponding to different constant values of magnetic field strength H, RF power P_k, and efficiency η. Figure 13-20 shows typical characteristics of a pulsed multicavity magnetron. These curves show that in order to increase the useful RF power, it is necessary to increase the magnetic field strength H and the anode voltage U_a, while in order to increase the efficiency η for a given anode current, it is necessary to increase the anode voltage. Conversely, an increase in anode current for a given anode voltage leads to a drop in η. Thus, the characteristics given make it possible to establish the desired operating regime for the magnetron. Thus, for example, for a useful RF power $P_k = 300$ kw and an efficiency $\eta = 45$ per cent, we will have an anode voltage of $U_a = 16$ kv and an anode current $I_a = 43$ amp (point P). In this case, the magnetic field strength should be on the order of 1700 oersteds.

13-5. TRAVELING-WAVE TUBES.

The first models of electron amplifier tubes designed for SHF service, which have come to be called traveling-wave tubes (abbreviated LBV*), were obtained in 1944.

At present, two types of traveling-wave tubes are produced: input types, possessing high sensitivity and low noise figure, and output tubes with output powers on the order of 1-20 watts. Input LBV are

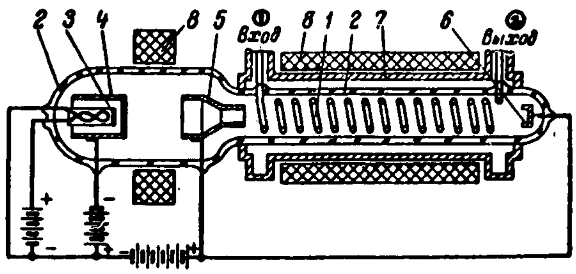

Fig. 13-21. Schematic diagram of traveling-wave tube and basic circuit used with it.
1) Input; 2) output.

utilized in radar receivers and in radio relay-link receivers. Output LBV are used in radio relay-link transmitters. The advantage of traveling-wave tubes in comparison with klystrons and magnetrons lies in the fact that they have a wide passband, as is needed for microwave work, and have a low internal-noise level, and are more efficient. The noise figure in a traveling-wave tube is less by a factor of 100 than for klystrons.

Traveling-wave tubes may oscillate in the 3000 Mc and above band.

The basic arrangement and circuit used with the traveling-wave tube is shown in Fig. 13-21. The basic element of the tube is a long hollow helix wound from wire, 1, inside an envelope 2, which has been evacuated. To one side of the helix there is a cathode 3, a control electrode 4, and an accelerating electrode 5; on the other side there is a collector 6, which is at the potential of the accelerating elec-

* [ЛБВ - LBV - lampa c begushchey volnoy - traveling-wave tube.]

trode. The spiral, together with the envelope, is located in a metal cylinder 7. The length of the spiral is on the order of 50 cm, and its diameter ranges from 1-3 mm.

The signal being amplified, which takes the form of an alternating current traveling along the wire of the spiral, as a wave, is applied to the end of the spiral located near the cathode, while the amplified signal is taken from the other end.

Electrons leaving the hot cathode are accelerated by the high positive potential on electrode 5, and progress along the axis of the helix at high velocities, reaching 0.1 times the speed of light. Upon leaving the spiral, the spent electrons are trapped by the collector. The electron beam is controlled by changing the potential on electrode 4, while the electrons are focused by an axial magnetic field set up by the coils 8.

In traveling-wave tubes, electrons moving along the axis of the helix interact continuously with the electrical field of the wave moving along the turns of the spiral; here the electrons give up a portion of their kinetic energy to the electromagnetic wave and amplify it.

As is known, effective interaction of electrons with a high-frequency field is achieved by two methods: either by increasing the force acting upon an electron (for example, in microwave triodes and in klystrons), or by increasing the path over which the force acts on the electron. The second method is utilized in traveling-wave tubes.

The signal applied to the input of the tube forms a wave which in a straight wire would be propagated in accordance with the fluctuations in the electrical charges at a velocity v_v, about equal to the speed of light $\underline{c}$. If the conductor is helical in form, the rate of movement of a wave along the turns of the helix remains unchanged,

v_v roughly equals c equals $3 \cdot 10^{10}$ cm/sec, while the speed of motion of the electric field formed within the spiral by the wave decreases; here the greater the diameter and the smaller the pitch of the helix, the lower the speed of movement of the electric field along the axis of the helix. The connection between the rate of motion of the wave v_v, the rate of motion of the axial electric field v_p,* the pitch p and the length of a single turn l = πd of the helix is defined by the relationship

$$\frac{v_\text{п}}{v_\text{в}} = \frac{p}{\pi d}.$$

From this we may find the velocity of the axial electric field along the axis of the helix.

$$v_\text{п} = v_\text{в} \frac{p}{\pi d} \approx c \frac{p}{\pi d}.$$

Selecting the ratio p/πd, the velocity of the axial field may be reduced to a value v_p = 0.1c, at which the axial field and the electrons of the beam will move at about the same velocity, interacting continuously along the entire path of motion within the helix.

The electromagnetic wave moving along the turns of the helix sets up within the helix not only an electric field, but a magnetic field as well. The action of the magnetic field on the stream of electrons may be neglected, however, since it is extremely small. Of interest from the practical point of view is the interaction of the electron beam with the electric field alone; within the helix, it is possible to distinguish three components: axial, radial, and tangential. Of these components, the radial and tangential equal zero on the axis of the helix (owing to symmetry), while near the axis they are so small

---

* [$v_\text{в}$ - v_v - v_{volna} - v_{wave}; $v_\text{п}$ - v_p - v_{pole} - v_{field}.]

that their effect upon the electron beam need not be taken into consideration. From this it follows that only the axial component of the traveling-wave electric field interacts with the stream of electrons.

Let us now consider the mechanism for electron interaction with the electric field in various sections of the traveling wave. Let a traveling wave move along the helix, which takes the form of a delay system, and at the input of the helix, for various portions of each wave, let electrons flow in, moving at the velocity of the wave. Clearly, those electrons that enter with portions of the wave where the electric field is zero will not change in kinetic energy, since along the entire path through the axis of the helix, they will move with the velocity of the wave, and at all times will be in sections where there is no field.

Those electrons that enter with a traveling-wave section where the electric field is accelerating for the electrons interact with the field (the field accelerates the electrons). As a result, the energy of the electrons is increased, while the energy of the traveling wave is decreased.

Finally, if electrons enter in a wave section where the field is retarding for the electrons, the electron kinetic energy will drop, while the traveling-wave energy will rise (the wave is amplified).

If the velocity of the electrons equals the velocity of the traveling wave, however, and if there is a uniform stream of electrons in the helix, the electrons will be distributed equally in accelerating and retarding sections of the wave. Thus, the mean energy of all electrons in the various sections of the wave, and the energy of the wave, will not change. Consequently, the wave as a whole will not be amplified. Clearly, for normal operation of a traveling-wave tube, it is necessary to provide conditions in which the electrons will be bunched chiefly in the retarding sections of the wave, while in the accelerating wave

sections, there will be as few as possible.

In order to clarify the conditions for such bunching, let us consider the interaction of an electron beam with a traveling wave.

1. Let the velocity of the electrons equal the velocity of the traveling waves. In this case, electrons in accelerating sections of the wave will increase in velocity, and will lose velocity, correspondingly, in retarding sections. The accelerated and delayed electrons gradually draw together and bunch in a section of the wave where the accelerating field replaces the retarding field. Since the number of electrons in accelerating sections of the wave equals the number of electrons in retarding sections of the wave, and retardation and acceleration occur to an equal extent, the energy of the wave will not be changed. Thus, a traveling-wave tube will not amplify if the electron velocity equals the traveling-wave velocity.

2. Let us now decrease the velocity of the electrons so that it becomes somewhat less than the velocity of the traveling wave. In this case, electrons falling into a wave region with an accelerating field, despite the fact that, according to the initial assumption, their velocity is somewhat less than the wave velocity, will remain in an accelerating field during their transit time through the spiral, since the accelerating field increases their velocity. Consequently, these electrons take energy from the traveling-wave field.

On the other hand, electrons falling into a section of the wave with a retarding field are retarded still more and fall still further behind the wave, gradually falling into a section of the wave with an accelerating field. Thus, if the velocity of the electrons is somewhat less than the velocity of the traveling wave, the electrons will be bunched in wave sections with accelerating fields. As a result, the energy of the electrons is increased, while the wave energy drops.

Consequently, LBV will not amplify in this case.

3. Let us now assume that the velocity of the electrons going into the traveling wave is somewhat higher than the velocity of the wave.

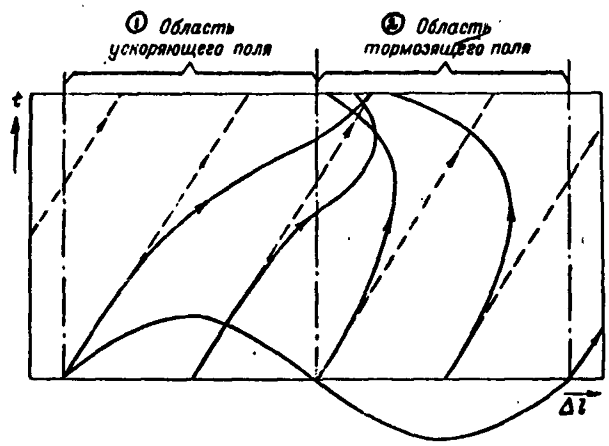

Fig. 13-22. Mechanism of electron bunching under the influence of a traveling wave (initial velocity of electrons somewhat less than the wave velocity). The dashed lines are electron trajectories in the absence of the traveling waves.
1) Accelerating-field region; 2) retarding-field region.

In this case, on sections of the wave with an accelerating field, the electrons will move somewhat faster than the wave (both as a result of the greater initial velocity, and as a result of their acceleration), while on sections with a retarding field, they will be somewhat slower than the waves, since they are retarded by the field. Thus, electrons will pass from the accelerating field into a retarding field in large numbers, while small numbers will pass from a retarding field into an accelerating field. As a result, the electrons are bunched on wave sections with retarding fields. In this case, a considerable portion of the kinetic energy will be given up by the bunched electrons to the field of the traveling wave, and as a result the wave and, consequently, the signal, will be amplified. Thus, for normal traveling-wave tube operation it is necessary for the velocity of the electrons

entering the wave to be somewhat higher than the wave velocity.

Figure 13-22 shows the change in electron velocity and the change in electron grouping thus created, corresponding to the last case; here, the shift in electrons relative to the wave $\Delta \underline{l}$ is plotted along the X axis, and time $\underline{t}$ along the Y axis.

The considerations presented are correct if there is only a traveling wave in the helix. In actuality, some energy may be reflected from the load (reflected wave). The reflected wave, superimposed upon the traveling wave, changes the field pattern in the helix. As a result, the energy-transfer mechanism becomes more complicated.

13-6. BACKWARD-WAVE TUBES.

In design and principle of action, a backward-wave tube is similar to a traveling-wave tube. The difference lies chiefly in the fact that in the backward-wave tube, use is made of the interaction of a beam of electrons with components of a complex electromagnetic wave that propagates in the opposite direction (opposite to the traveling wave). From this is derived the name "backward-wave tube" or, as abbreviated, LOV*.

In a backward-wave tube, the energy of the beam of electrons and the energy of the backward wave propagating in the opposite direction interact for many oscillation cycles.

It has been established that with a backward wave, both amplification and generation of oscillations take place. The peculiarity of the backward-wave tube lies in the fact that it may be easily converted from oscillator to amplifier by changing only a single accelerating voltage (electronic tuning). This voltage normally varies from several hundreds to several thousands of volts. In this case, the frequency of

*[ЛОВ - LOV - lampa c obratnoy volnoy - backward-wave tube.]

the oscillations generated varies over a wide range reaching about 60,000 Mc. A backward-wave oscillator tube gives good frequency stability, which, as far as is known, is independent of the load, while its efficiency is comparable with the efficiency of a reflex klystron.

In a backward-wave oscillating tube designed to operate at very high frequencies, instead of the normal delay system, for example a helix, another system is used in which there is a series of slots located a specific distance apart, and in which the beam of electrons is shielded from the wave over a considerable segment of its path. In this system, the beam of electrons interacts with the electromagnetic waves only in the slots. Thus, in a tube of this type, the velocity of the electromagnetic wave does not decrease to the velocity of the electron beam. In LOV, the interaction of the electrons with the wave is similar to the interaction taking place in LBV, but is different.

Chapter Fourteen

PHOTOELECTRIC DEVICES

14-1. BASIC CONCEPTS OF THE QUANTUM THEORY OF LIGHT.

We established in Chapter One that according to Planck's theory, radiant energy may be emitted and absorbed in definite portions (quanta) whose magnitude is determined by the product $h\nu$. In 1905, Einstein developed this theory further, proposing that if radiant energy is emitted and absorbed in quanta, the radiation process itself is quantized, i.e., the propagation of luminous energy is not wave propagation, but the propagation of a kind of particle (corpuscle) known as a light quantum or photon. Thus, in accordance with the Planck-Einstein conceptions, light not only leaves the atom in the form of an energy quantum, but is also propagated in space and enters into interaction with matter in the form of the same energy quantum $h\nu$, which travels through space as a whole with the velocity of light.

Photons do not exist in a state of rest, and consequently have no resting mass; in motion, however, they possess a definite mass m which depends on the oscillation frequency ν. This relation is given by the equation $m = h\nu/c$, where c is the velocity of light.

Upon striking solids, photons transfer energy to them and create the so-called light pressure, which was first experimentally observed by P.N. Lebedev in 1900.

When a photon is absorbed by a material body, its energy is used to increase the energy of one of the electrons of an atom, and can be absorbed only as a whole in this process. If the photon's energy is less than the energy required to promote an electron to a higher energy level, the photon is scattered without a change in its energy. If the photon's energy is sufficiently large, however, the electron may jump

to a higher energy level or break away from the atom on absorbing it. The so-called photoionization, as a result of which a positive ion and a free electron form and the photon disappears after giving up its energy, occurs in the latter case. On the other hand, if the electron passes from a higher to a lower energy level, the atom emits a quantum of luminous energy (a photon) into space.

Thus, according to the quantum theory, the emission and absorption of luminous energy by matter are discontinuous (discrete) in nature, i.e., can take place only in finite portions (quanta, photons). Here, this emission also possesses wave properties together with its corpuscular properties, by virtue of which a definite frequency or wavelength is assigned to each quantum. This is a manifestation of the dual nature of light. This duality in the science of light is upheld by quantum mechanics, which, in studying the laws of movement and interaction of particles of atomic dimensions which possess little mass, has established their corpuscular and wave properties as well as the discontinuity of their interaction.

14-2. CLASSIFICATION OF PHOTOELECTRONIC DEVICES.

The phenomena which unfold in the interaction of light, or radiant energy in general, with bodies and are accompanied by the transformation of luminous (radiant) into electrical energy are commonly called the photoelectric effect or, briefly, the photoeffect. Three forms of the photoeffect are known at the present time:

1. The extrinsic photoeffect for photoelectron emission, in which electrons are torn from the surfaces of bodies under the influence of light.

2. The intrinsic photoeffect, in which the conductivity of semiconductors and dielectrics changes under the influence of light due to

the appearance of additional free electrons (conduction electrons) inside the semiconductor.

3. The photovoltaic effect, in which an emf arises under the influence of light falling on the boundary between a metal and a light-sensitive semiconductor, leading to the appearance of, or a change in, a current in an external circuit.

Devices in which the photoelectric effect is utilized have come to be known as photocells. In accordance with the forms of the photoeffect listed above, photocells are divided into three basic groups: 1) photocells with extrinsic photoeffect; 2) photocells with intrinsic photoeffect (photoresistors); 3) photocells with photovoltaic effect (photovoltaic cells). Photocells with extrinsic photoeffect are, in turn, classed as vacuum and gas-filled types (the latter are considered in the course on ionic devices; photoresistors and photovoltaic cells are treated in Chapter Seventeen).

In addition to photocells, there are great numbers of more complex devices in which other phenomena are used together with the photoeffect. These include: 1) multiplier tubes (FEU*), in which photoelectronic and secondary-electron emission and electron-optical phenomena are used; 2) electron-optical transducers (EOP**), in which photoelectronic emission, electron-optical phenomena, and luminescence phenomena are used; 3) iconoscopes, in which photoelectron emission and electron-optical phenomena are used (electron-optical transducers and iconoscopes are considered in the following chapter).

The above types of photoelectronic devices are, in turn, subdivided in accordance with their design makeup or in accordance with the type of

---

* [ФЭУ - FEU - Fotoelektronnyy Umnozhitel' - Photomultiplier.]

**[ЭОП - EOP - Elektronno-Opticheskiy Preobrazovatel' - Electron-Optical Transducer.]

photocathode used.

14-3. PHOTOCATHODES.

Photocells with intrinsic photoeffect differ from other types of photocells in possessing light-sensitive cathodes (photocathodes) which emit streams of free electrons when acted upon by light. Simple metallic cathodes have limited applications in photocells because they do not provide sufficiently strong electron emission. They are used, for example, in the manufacture of devices which are sensitive in the ultraviolet region but virtually insensitive in the visible region of the spectrum. The so-called composite photocathodes, which have a complex-structured photosensitive layer formed either on the inside wall of the glass bulb or on a special metallic electrode (base layer) inside the bulb, possess considerably higher photoelectronic emissivities.

The most widely used composite photocathodes are the oxygen-cesium, antimony-cesium, bismuth-silver-cesium, and multi-alkali photocathodes.

1. Oxygen-cesium photocathodes. Cathodes of this type, which are also known as oxygen-silver-cesium cathodes, have a silver base layer, which is deposited on part of the inside surface of the glass bulb and oxidized by electric discharge in oxygen. In order to sensitize the layer, cesium vapor is introduced into the tube and reduces the silver oxide on contact with the oxidized film, thus forming a cesium-oxide film; some of the cesium is deposited in pure form on the cesium-oxide layer. Such a photocathode has its highest photoelectron emission when a monatomic cesium layer has been formed on its surface. On the other hand, however, as the thickness of this layer increases, the work function increases and the photoelectron emission falls off.

The structure of the oxygen-cesium photocathode is represented

schematically in Fig. 14-1. It should be noted that a well-activated cathode should not have an Ag_2O interlayer.

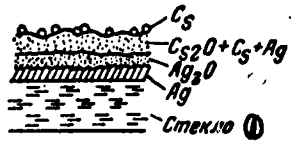

Fig. 14-1. Structure of oxygen-cesium photocathode. 1) Glass.

Oxygen-cesium photocathodes are utilized in photocells of types TsV-1, TsV-3, and TsV-6.*

2. Antimony-cesium photocathodes. In the manufacture of antimony-cesium photocathodes, the antimony is deposited on part of the inside surface of the glass bulb by evaporation from a nickel or tungsten wire heated by an electric current. The thin film of antimony that forms on the glass is treated with cesium vapor until the photocurrent maximum is obtained. In order to increase the photoelectron emission, the cathode is usually lightly oxidized. Marked oxidation, however, results in decreased emission.

Antimony-cesium photocathodes have a number of advantages over oxygen-cesium photocathodes. First of all, they possess high quantum efficiencies, running up to 25 per cent; the quantum efficiency of an oxygen-cesium photocathode reaches only 2-3 per cent, while that of purely metallic photocathodes is approximately 0.1 per cent. In addition, antimony-cesium photocathodes are more resistant to temperature fluctuations than oxygen-cesium photocathodes.

Antimony-cesium cathodes are utilized in photocells of types STsV-3, STsV-4, STsV-51, etc. (STsV stands for antimony [Sur'ma]-cesium vacuum).

* [ЦВ - TsV - tseziyeviy, vakuumnyy - cesium, vacuum.]

3. Bismuth-silver-cesium photocathodes. In the manufacture of cathodes of this type, a thin layer of bismuth is applied to a region on the inside surface of the glass bulb by spraying and subsequently oxidized by electric discharge in oxygen, with an increase in transparency. Then silver is sprayed on over the lightly oxidized bismuth layer. The layer of silver formed over the oxidized bismuth on the inside surface of the bulb is subjected to further treatment with cesium vapor. The cesium dose is adjusted during the process of sensitizing by appropriate heat treatment of the cathode until the photocurrent maximum is reached; in this process, excess cesium is removed.

As compared to antimony-cesium photocathodes, bismuth-silver-cesium photocathodes are light-sensitive not only to those wavelengths to which the antimony-cesium photocathodes respond (4,000-7,000 A), but also to longer wavelengths (from 4,000 to 8,000 A), although their sensitivity in the region from 4,000 to 6,000 A is somewhat lower than that of the antimony-cesium photocathode.

4. Multi-alkali photocathodes. Multi-alkali photocathodes are prepared by appropriate treatment of a thin layer of antimony with potassium, sodium and cesium vapors. The basic advantage of these photocathodes over the composite photocathodes described earlier is their high sensitivity (100-190 μa/lumen) in the long-wave region of the spectrum, coupled with relatively low dark currents (multi-alkali effect). The spectral-response region of the multi-alkali photocathode is 4,000-8,500 A.

14-4. PHOTOELEMENTS WITH EXTRINSIC PHOTOEFFECT.

1. Physical design.

The basic physical design of a photocell with extrinsic photoeffect is shown schematically in Fig. 14-2. The photocathode k is part

of the inside surface of the glass bulb b to which a light-sensitive layer has been applied. A metallic lead c is run through the wall of the bulb from the photocathode. The part of the bulb opposite the photocathode is transparent and serves as the window (optical input) of the photocell, through which the luminous flux falls on the cathode. A plate a, the distance and shape of which are so selected that it will interfere only slightly with the incidence of the luminous flux onto the photocathode and at the same time act as a sufficiently good electron collector, is located inside the bulb. Usually, the plate takes the form of a ring, a screen, or a frame and is located either in the center of the tube or slightly below it.

If the photocathode is connected to the negative pole of the source and the plate to the positive pole, the cathode will begin to emit electrons when it is illuminated in the presence of a high vacuum; moving toward the plate in the accelerating field, these electrons set up in the external circuit a weak electric current which varies in proportion to the luminous flux. This current can be detected by the sensitive galvanometer G connected into the circuit.

The basic element of the photocell with extrinsic photoeffect is the photocathode, the quality of which predetermines the properties of the photocell.

In photocells with oxygen-cesium photocathodes, the plate is most often made in the form of a small ring located symmetrically with respect to the cathode at the center of the bulb, or in the form of a frame which surrounds the cathode. In this design, the plate does not interfere with the incidence of light onto the photocathode.

In order to increase the insulation resistance and thereby decrease the leakage currents across the outside surface of the bulb, some types of photocells are given only one support, to which the lead from the

cathode or the plate is connected.

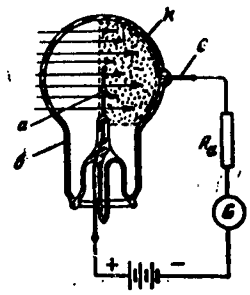

Fig. 14-2. Schematic diagram of physical design and wiring of photoelement with extrinsic photoeffect.
a – Plate; b – glass bulb; c – cathode lead; G – galvanometer; k – cathode; R_a – load resistance.

In photoelements with antimony-cesium photocathodes, the plate need not be located in the center, since the antimony is oxidized not by electrical discharge in oxygen, but by gradual treatment of the $SbCs_3$ layer with small portions of oxygen during continuous evacuation of the photoelement. In order not to prevent the light passing through the photocell window from striking the cathode, the plate of this type of photocell is, as a rule, shifted downward.

In addition to the above types of photocells, there are other "double-plate" photocells which differ from the normal type in possessing a second plate. Such photocells are designed for work in special bridge circuits. The double-plate STsV-2A photocell with antimony-cesium cathode may serve as an example of this type of photocell.

In order to increase the sensitivity to light of shorter wavelengths, which is absorbed by ordinary glass, some photocell types are made with special (uviol) glass, which passes light of shorter wavelengths. The phototube window is made thinner for this purpose. In some cases, a window of thin uviol glass is sealed into a bulb of ordinary

glass. The type STsV-6 photoelement may serve as an example of a uviol-bulb photoelement.

2. Characteristics and parameters.

The proportionality factor k which appears in Eq. (14-3) is known as the sensitivity or integral efficiency of the photocell and characterizes the ability of the photocell to react to a luminous flux of the full spectral composition (white light). However, photocathodes respond differently to oscillations of different frequencies. For this reason, in addition to the integral efficiency, i.e., the photosensitivity to the entire spectrum from the ultraviolet to the infrared, we distinguish the sensitivities of a photocell to light of one and only one wavelength (so-called monochromatic radiation), which is known as its spectral sensitivity.

The differences in the sensitivities of photocells to radiation of different wavelengths may be accounted for first, by the fact that the energy of the quantum increases with increasing frequency of the radiation, so that electrons which possessed small energies before acquiring the energy of a quantum may leave the cathode, while close to the photoelectric threshold only those electrons which had energies close to the maximum W_i can leave. For this reason, sensitivity generally decreases in all types of photocathodes with increasing wavelength of the incident light. Secondly, certain photocathodes may possess one or several sensitivity maxima, even at not particularly high frequencies. This is because these photocathodes have energy states of their electrons such that the transfer probability for the energy of the quanta is greatest at the specific energy values of these quanta. As a result, one and the same photocathode, not to mention different types of photocathodes, gives one photoelectron-emission value when illuminated by light of one wavelength and another when it is illuminated by light of

another wavelength.

The diagram of spectral sensitivity as a function of the wavelength or frequency of the light incident on the cathode is usually known as the spectral characteristic or spectral-sensitivity distribution of the photocathode. Typical spectral sensitivity distributions of oxygen-cesium and antimony-cesium photocathodes are shown in Fig. 14-3, where the values of the wavelengths of light in Angstrom units are plotted against the X axis and the ratio of the sensitivity at a given wavelength to the maximum sensitivity in the visible or near-infrared region of the spectrum is plotted against the Y axis as a function of the type of photocathode used in the photocell. These curves show that photocells with antimony-cesium photocathodes possess their highest sensitivities at $\lambda \approx$ 4,500 A, i.e., in the green and blue regions. However, at wavelengths of 6,500 A and higher, i.e., in the red region, photocells with antimony-cesium photocathodes have, for all practical purposes, zero sensitivities. On the other hand, photocells with oxygen-cesium photocathodes possess their highest sensitivities at $\lambda \approx$ 7,500-8,500 A, i.e., in the red region, and at $\lambda \approx$ 3,500 A, i.e., in the violet region. Thus, practical utilization of any photocell requires knowledge of its spectral characteristic and the use of light of that spectral composition which gives rise to the greatest photoelectron emission.

The spectral characteristic of a photocell with an antimony-cesium photocathode in a uviol glass bulb is shown in Fig. 14-4. This characteristic shows that a uviol glass bulb permits extending the spectral sensitivity region of the antimony-cesium photocell from 4,000 to approximately 2,000 A.

Apart from the spectral characteristic, the so-called light characteristic of a photocell, which expresses the photoelectron emission as a function of the luminous-flux intensity at a steady plate voltage

the magnitude of which must correspond to the saturation mode, is of great practical significance. It follows from the equation i_a* = kF that the light characteristic is a straight line proceeding from the coordinate origin. Such characteristics are shown in Fig. 14-5 for oxygen-cesium and antimony-cesium photocells.

The equation $i_a = kF$ is valid for photoelectron emission, but the linearity of the light characteristic is disturbed when other types of emission occur. Thus, for example, antimony-cesium (with metallic base layer) and oxygen-cesium photocells have linear light characteristics, while antimony-cesium photocells without metallic cathode base layers have light characteristics whose linearity is maintained only for small luminous-flux values; for large luminous fluxes, however, linearity is violated (Fig. 14-6). This is accounted for by the fact that an

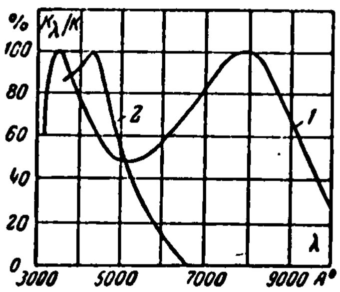

Fig. 14-3. Spectral characteristic of oxygen-cesium (1) and antimony-cesium (2) photocells.

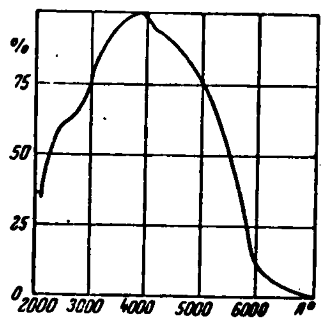

Fig. 14-4. Spectral characteristic of STsV-6 photocell in uviol glass bulb.

antimony-cesium photocathode without a metallic base layer has a high resistance, and when strong photoelectron emission currents due to high luminous fluxes pass through the cathode, a large voltage drop is set up between regions of the photocathode adjacent to the photocathode lead-in and regions remote from the lead-in. Consequently, regions remote from the cathode are at higher positive potentials than are regions

* [i_a - i_a - i_{anod} - i_{plate}.]

adjacent to the lead. The farther a region is from the attachment of the lead, the higher will its positive potential be. For this reason, a certain fraction of the electrons leaving regions adjacent to the lead

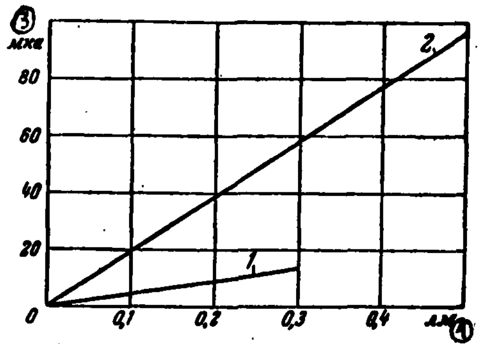

Fig. 14-5. Light characteristics of vacuum photocells.
1) Oxygen-cesium; 2) antimony-cesium; 3) μa; 4) lumens.

in response to illumination may move not to the plate, but to regions of the cathode distant from the lead and produce secondary-electron emission from these regions by their bombardment. Consequently, the

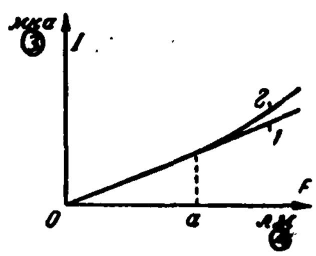

Fig. 14-6. Light characteristics.
1) Antimony-cesium photocell with metallic base layer, and oxygen-cesium photocell; 2) antimony-cesium photocell without metallic base layer; 3) μa; 4) lumens.

plate current will rise in this case not only as a result of increased photoelectron emission brought about by increased luminous flux, but also as a result of the ever-increasing secondary-electron emission. As a result, the linearity of the light characteristic will be disturbed at high luminous fluxes. Deviations from linearity may also be brought about by the presence of a space charge or by ionization of

residual gases.

Linearity of the light characteristic ensures undistorted reproduction of the variation of the luminous flux in the variation of the electric current. For this reason, photocells with linear characteristics are used for undistorted reproduction of the luminous-flux variation, or we work with luminous-flux values corresponding to the linear region of light characteristic (for example, the region 0a of Fig. 14-6).

The light-characteristic nonlinearity sometimes observed at high luminous fluxes in vacuum photocells with oxygen-cesium photocathodes is accounted for by the fatigue effect, which consists in a decline in the sensitivity of the photocell in the course of operation (especially when it is under continuous illumination). If, however, the photocell's cathode is made of pure, thoroughly outgassed metal, low fatigue effect is observed.

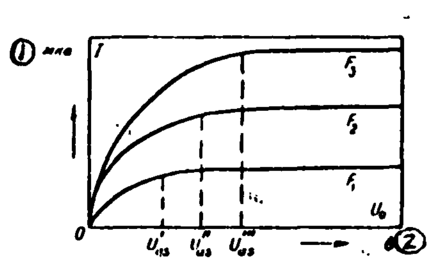

Fig. 14-7. Volt-ampere characteristic of vacuum photocell at different luminous-flux values ($F_3 > F_2 > F_1$). 1) μa; 2) volts.

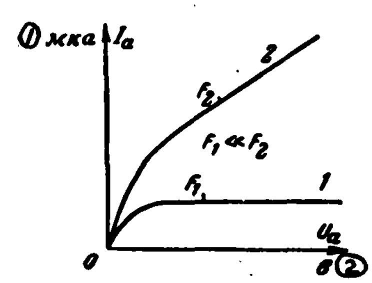

Fig. 14-8. Volt-ampere characteristic of antimony-cesium vacuum photocell without metallic base layer for different luminous-flux values. 1) μa; 2) volts.

To select the operating mode, it is necessary to know the volt-ampere characteristic of the photocell, which is a diagram of the photoelectron current as a function of plate voltage at constant luminous flux. The volt-ampere characteristics of antimony-cesium vacuum-type photocells (with metallic cathode base layers) and oxygen-cesium photo-

cells, which express the photoelectron current qualitatively as a function of plate voltage, are shown in their general form in Fig. 14-7 for three luminous-flux values. The characteristics show that the photocurrent first increases with increasing plate voltage, but then moves into a saturation region and does not change with further increases in plate voltage. The magnitude of the saturation current depends on that of the luminous flux. The larger the luminous flux, the higher will be the saturation current, and the higher the plate voltage at which it will stabilize. The magnitude of the saturation plate voltage depends, in turn, on the design of the photocell and on the intensity of the illumination incident upon the photocathode surface.

As the volt-ampere characteristics show, the operating voltages of vacuum photocells can be varied over wide ranges corresponding to the saturation region, especially when they are working at low luminous fluxes.

At small luminous fluxes, when secondary electron emission from the remote regions of the cathode has not yet begun, the volt-ampere characteristics of antimony-cesium photocells without metallic cathode base layers take the same form as the characteristics described above for photocells with metallic cathode base layers (curve 1, Fig. 14-8).

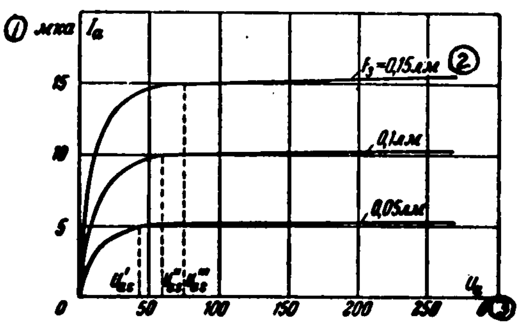

Fig. 14-9. Volt-ampere characteristics of types STsV-3, STsV-4, and STsV-51 antimony-cesium vacuum photocells.
1) μa; 2) lumens; 3) volts.

At high luminous fluxes, however, the photoelectron-emission current increases continuously with plate voltage because of secondary-electron emission from remote regions of the cathode and does not enter a saturation region (curve 2, Fig. 14-8).

The volt-ampere characteristics of types STsV-3, STsV-4, and STsV-51 antimony-cesium photocells, which express plate current qualitatively and quantitatively as a function of plate voltage, are presented in Fig. 14-9 for three different luminous-flux values. These characteristics are typical for all three photocell types.

Since the plate current follows luminous-flux variations almost instantly as a result of the small mass of the electrons, vacuum photocells are fast-responding devices. It has been established experimentally that photocurrent changes follow no later than 10^{-9} sec after luminous-flux changes in vacuum phototubes. For this reason, the sensitivity (or photocurrent) of vacuum photocells of all types remains the same for fast luminous-flux variations as at steady luminous fluxes, i.e., is virtually independent of the frequency of variation (modulation) of the luminous flux.

Knowledge of their parameters is sufficient for determination of the operating-mode properties of photocells; these include the minimum initial sensitivity, the maximum dark current, which flows in the cell in total darkness, and the operating voltage. The sensitivities after continuous operation over their service lives are also given for certain types of photocells.

The basic parameters of vacuum photocells are given in Table 14-1.

In measuring the parameters of vacuum photocells with extrinsic photoeffect, the nominal voltage (240 v) is applied to the photocell. An incandescent lamp with a tungsten filament and a color temperature of $2,848^{\circ}$ K is used as the light source. A load resistance R = 200 kohm

TABLE 14-1.
Basic parameters of vacuum photocells.

1 Типы фотоэлементов	2 Рабочее напряжение, в	3 Минимальная начальная чувствительность, мка/лм	4 Спектральная область, Å	5 Максимальное значение темнового тока, а	6 Минимальное среднее значение чувствительности после непрерывной работы в течение 1 000 ч, мка/лм
7 ЦВ-1	240	20	6 000—11 000	$1 \cdot 10^{-7}$	Не нормируется 14
8 ЦВ-2	240	20	6 000—11 000	$1 \cdot 10^{-7}$	» »
9 ЦВ-3	240	20	6 000—11 000	$1 \cdot 10^{-7}$	» »
10 СЦВ-3	240	80	3 800—7 000	$1 \cdot 10^{-8}$	60 при 0,005 лм 15
11 СЦВ-4	240	80	3 800—7 000	$1 \cdot 10^{-7}$	60 при 0,01 лм 16
12 СЦВ-51	240	80	3 800—7 000	$1 \cdot 10^{-8}$	60 при 0,05 лм 17
13 СЦВ-6	30	60	2 150—7 000	$5 \cdot 10^{-11}$	Не нормируется 14

1) Type of photocell; 2) operating voltage, v; 3) minimum initial sensitivity, μa/lumen; 4) spectral region, A; 5) maximum dark current, amp; 6) minimum average sensitivity after 1000 hours of continuous operation, μa/lumen; 7) TsV-1; 8) TsV-2; 9) TsV-3; 10) STsV-3; 11) STsV-4; 12) STsV-51; 13) STsV-6; 14) not standardized; 15) 60 at 0.005 lumen; 16) 60 at 0.01 lumen; 17) 60 at 0.05 lumen.

is connected into the photocell circuit. The magnitude of the luminous flux incident on the cathode of the photocell must be between 0.002 and 0.05 lumen, while the area of the circular diaphragm aperture is made to conform with the data of Table 14-2.

The photocell must be completely darkened for dark-current measurements.

A serious photocell defect is the so-called noise which arises when the current in the photocell circuit deviates from an average value, i.e., on current fluctuations brought about, for example, by fluctuations in photoelectron emission (shot effect), fluctuations in the light source, thermal fluctuations in the load resistance, and by changes in the contact resistance between the lead-in and the photocathode. The noise is most pronounced at high photocurrent gains, when the fluctuations are amplified together with the photocurrent and become audible as interference at the amplifier output. Admittedly, this

noise is virtually imperceptible in vacuum photocells when the photocurrent amplification is small.

TABLE 14-2.

① Типы фотоэлементов	② Площадь круглого отверстия диафрагмы, $см^2$
③ ЦВ-1	10
④ ЦВ-4	5
⑤ СЦВ-4	5
⑥ СЦВ-3	2,5
⑦ ЦВ-3	2,5
⑧ СЦВ-51	3,14

1) Photocell type; 2) area of round diaphragm aperture, cm^2; 3) TsV-1; 4) TsV-4; 5) STsV-4; 6) STsV-3; 7) TsV-3; 8) STsV-51.

Moreover, the photoelectron emission declines noticeably in the course of prolonged operation irrespective of the photocathode type used, and is restored only partially after operation is suspended (fatigue effect). Full restoration of photoelectron emission usually does not occur.

The fatigue effect is considerably less pronounced in photocells with antimony-cesium photocathodes than in photocells with oxygen-cesium photocathodes (Fig. 14-13).

14-5. PHOTOMULTIPLIERS, THEIR PARAMETERS AND CHARACTERISTICS.

The sensitivities of the extrinsic-photoeffect photocells considered above are very low, averaging a few tens of microamperes per lumen. In practice, however, luminous fluxes running to tenths, hundredths, or even thousandths of a lumen are frequently used. The operating currents of the photocells naturally amount to microamperes or fractions of a microampere at such small luminous fluxes. Practical utilization of such small currents requires preamplification by means of rather cumbersome multiple-tube amplifier circuits.

In 1934, L.A. Kubetskiy suggested that the photocurrent be multiplied not by a separate tube amplifier, but inside the photocell itself, utilizing the phenomenon of secondary electron emission for this purpose. Such photocells have come to be known as photomultipliers.

At the present time, the industry is producing two types of photomultipliers: single-stage designs intended for comparatively small photocurrent amplifications, and multiple-stage devices which produce considerable amplification of the photocurrent.

A. Single-stage photomultipliers.

The single-stage photomultiplier is a vacuum phototube into which an additional electrode has been introduced. To distinguish it from the photocathode, this electrode is conventionally known as the emitter. The photocathodes and emitters of types FEU-1 and FEU-2 single-stage photomultipliers are manufactured in the same way and consequently have identical properties. As regards their functions, however, they differ essentially: the cathode serves as the photoelectron (primary-electron) source, while the emitter is a secondary electron source.

The basic working principle and a schematic circuit diagram of a single-stage photomultiplier are given in Fig. 14-10. Some of the primary electrons emitted by the photocathode k under the influence of the luminous flux F are accelerated by the electric field and strike the plate a, creating in its circuit a primary-electron current I_1 which can be measured by the galvanometer G_1. Other accelerated primary electrons pass through the screen-type plate, striking the emitter e at high velocity and dislodging secondary electrons from it. The secondary electrons that leave the emitter are accelerated by the plate's electric field and, striking the plate, set up a current I_2 (secondary-electron current) in its circuit. This current can be measured with the galvanometer G_2. The galvanometer G in the plate circuit measures a

total current which is equal to the sum of the currents I_1 and I_2 and which exceeds in magnitude the total primary electron current. Thus, a phototube with an additional electrode (the emitter) amplifies ("multiplies") the photocurrent in and by itself. The term "photomultiplier" grew out of this circumstance.

The general appearance of the two single-stage photomultiplier types FEU-1 and FEU-2 appear in Fig. 14-11.

The designs of the FEU-1 and FEU-2 multipliers are similar to those of the corresponding photocell types. The cathode and the emitter are applied in the form of sensitized antimony-cesium layers to the inside surface of the glass bulb at points approximately opposite one another; the plate is located between them and, as in the antimony-cesium photocell, is shifted a short distance downward.

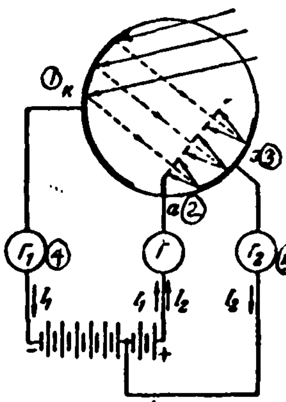

Fig. 14-10. Schematic circuit diagram of single-stage photomultiplier.
1) Cathode; 2) plate; 3) emitter; 4) galvanometer G_1; 5) galvanometer G_2; 6) galvanometer G.

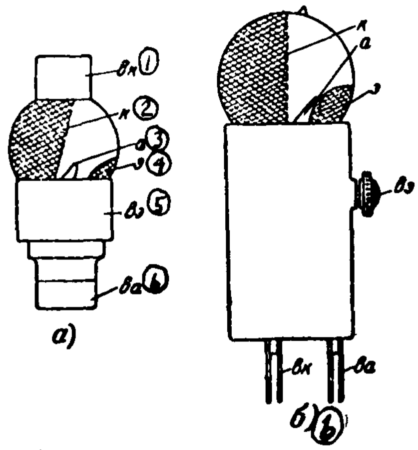

Fig. 14-11. General appearance of single-stage photomultipliers. a) FEU-2; b) FEU-1.
1) Cathode lead-out; 2) cathode; 3) plate; 4) emitter; 5) emitter lead-out; 6) plate lead-out.

The characteristics of single-stage photomultipliers are basically determined by the properties of the photocathodes utilized in them. Thus, for example, the spectral characteristics of the FEU-1 and FEU-2

multipliers, which have antimony-cesium photocathodes, are similar to the spectral characteristics of photocells with antimony-cesium photocathodes.

In addition to the spectral characteristics, the volt-ampere and light characteristics of photomultipliers are also of great practical importance. The total volt-ampere characteristic for the FEU-1 and FEU-2, which expresses the total luminous sensitivity (or the photocurrent corresponding to it) as a function of the multiplier's full supply voltage, is given in Fig. 14-12. This characteristic shows that as the total source voltage is increased, the photoelectron emission current rises rapidly at first and then more and more slowly. At volt-

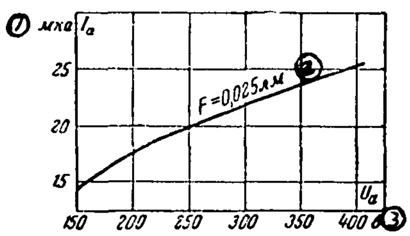

Fig. 14-12. Total volt-ampere characteristic of FEU-1 and FEU-2 single-stage photomultipliers.
1) μa; 2) lumens; 3) volts.

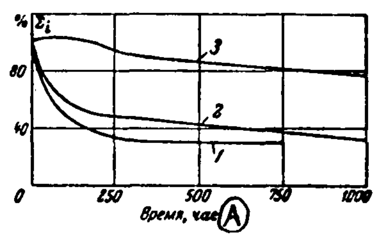

Fig. 14-13. Average fatigue curves of photocells and single-stage FEU photomultipliers for continuous operation.
1) For oxygen-cesium photocell; 2) for single-stage photomultiplier with antimony-cesium cathode; 3) for antimony-cesium photocell.
A) Time, hours.

ages higher than 250 v, the current increment is so small that the curve becomes approximately linear. This is accounted for by the fact that the velocities of the primary electrons increase considerably with increasing voltage. Some of these electrons have the maximum velocity and penetrate more deeply into the emitter, but, as we saw in Chapter Three, the secondary electrons dislodged by them from great depths lose velocity or change direction on the way toward the surface and hence do

not reach it. For this reason, the secondary-electron current begins to increase more slowly than at low plate voltages.

The light characteristic is linear in the range of luminous fluxes from 0.005 to 0.02 lumen, with the multiplier's photocathode illuminated by an incandescent electric lamp whose filament has a color temperature of 2,848° K.

Finally, to determine the properties of the FEU-1 and FEU-2 multipliers, it is helpful to know the curves of sensitivity variation of the multipliers during continuous operation. Figure 14-13 presents for comparison the average sensitivity curves of the FEU-1 and FEU-2 multipliers and photocells with antimony-cesium and oxygen-cesium photocathodes as functions of the time for which they have been operated. These curves show that the sensitivity of oxygen-cesium photocells drops sharply during their operation. The sensitivity of photocells with antimony-cesium photocathodes shows a much smaller decrease. However, despite the fact that they have antimony-cesium photocathodes, the sensitivity of the multipliers decreases much more noticeably than do the sensitivities of photocells with antimony-cesium photocathodes. This is accounted for by the fact that during operation, the multiplier emitters are broken down more rapidly than the photocathode. Nevertheless, the sensitivity loss of the multipliers during operation is somewhat smaller than that of photocells with oxygen-cesium photocathodes.

When photomultipliers (or photocells) are operated intermittently, their sensitivities are partially (or completely) restored during the interruptions in their work. In subsequent operation, however, they may drop to the previous values more quickly than at the beginning of operation.

This loss of sensitivity by photomultipliers and photocells during operation has come to be known as photocell or multiplier fatigue. The

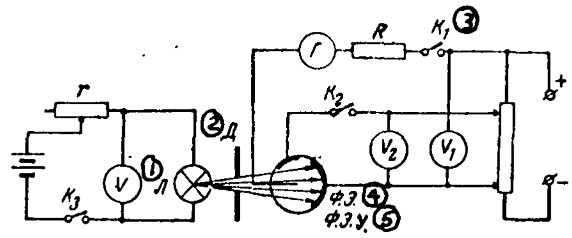

Fig. 14-14. Schematic circuit diagram for study of photocells and single-stage FEU (the multipliers are studied with K_1 and K_2 closed and the photocells with K_1 closed and K_2 open).
1) Lamp; 2) diaphragm; 3) switch; 4) FE; 5) FEU.

fatigue effect should not be confused with the phenomenon of photocell (or photomultiplier) aging, which is a sensitivity loss that takes place in photocells (multipliers) during storage, whether or not they are illuminated during this period.

The circuit shown in Fig. 14-14 is used to determine the parameters and characteristics of photocells and single-stage photomultipliers. In recording the characteristic of a photomultiplier and measuring its parameters, only the total source-voltage changes, while the distribution of the voltage among the electrodes remains the same.

TABLE 14-3.
Parameters of type FEU-1 and FEU-2 single-stage antimony-cesium photomultipliers.

1 Типы умножителей	2 Номинальное напряжение между анодом и катодом, в	3 Номинальное напряжение между эмиттером и катодом, в	4 Наименьшая интегральная чувствительность, мка/лм	5 Наибольший темновой ток на выходе умножителя, а
6 ФЭУ-1	220	170	400	$1 \cdot 10^{-7}$
7 ФЭУ-2	220	170	400	$1 \cdot 10^{-7}$

1) Multiplier type; 2) nominal voltage between plate and cathode, v; 3) nominal voltage between emitter and cathode, v; 4) minimum total sensitivity, μa/lumen; 5) maximum dark current at multiplier output, amp; 6) FEU-1; 7) FEU-2.

The basic parameters of the FEU-1 and FEU-2 photomultipliers are

shown in Table 14-3.

The tabulated data and the general appearance of the photomultipliers show that the FEU-1 and FEU-2 multipliers differ only in external design and have no basic differences as regards their construction and properties (parameters and characteristics) (Fig. 14-11).

B. Multiple-stage photomultipliers.

The schematic circuit diagram of a multiple-stage photomultiplier appears in Fig. 14-15. Several electrodes are placed inside the glass bulb from which the air has been eliminated: these are the photocathode k, which emits primary electrons when acted upon by a luminous flux, the plate a, and several emitters e, which are located equidistant from one another in a definite order between the cathode and the plate and are secondary-electron sources when bombarded by the primary electrons.

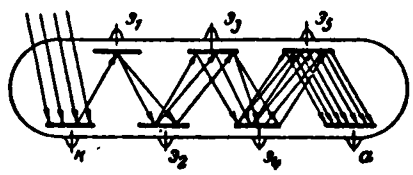

Fig. 14-15. Schematic circuit diagram of multiple-stage photomultiplier.
e_1 to e_5 – emitters; a – plate; k – cathode.

The number of emitters may vary, but it rarely exceeds 20 in practice. As in single-stage multipliers, the photocathode and emitters in multiple-stage multipliers may be made with identical or different properties. In practice, emitters of different types are used: oxygen-cesium, antimony-cesium, oxygen-magnesium, oxygen-beryllium, etc. Since the emitters (particularly those closer to the plate) are under greater current loads during operation of the multipliers than is the photocathode, their sensitive layers are made somewhat thicker than the sensitive layer on the photocathode.

The photocathodes used in photomultipliers are the same as those used in photocells, i.e., are usually of the antimony-cesium and oxygen-cesium type.

If a positive potential with respect to the cathode, for example, $U_1 = 150$ v is applied to the first emitter, and a potential $U_2 = 300$ v to the second emitter, etc., an accelerating electric field will arise between the cathode and the first emitter, between the first and second emitters, etc., for electrons moving from the cathode (or the emitters) toward the plate. The potential difference between each successive pair of emitters is usually 100 to 200 volts. Acted upon by a luminous flux, the photocathode emits primary electrons, which are accelerated by the electric field, strike the first emitter e_1 at high velocity, and dislodge a large number of secondary electrons from it. Under the influence of the accelerating electric field, these secondary electrons hit the second emitter e_2 with respect to which they are now primary electrons, and dislodge from it a larger number of secondary electrons, and so forth. Thus, the electrons pass successively from the cathode to the first emitter, from the first emitter to the second, etc., with their number growing during the process. Then the secondary electrons dislodged from the last emitter strike the plate a, which has a higher potential, and set up in the external circuit a current which exceeds the primary photocurrent i_0 created by the cathode under the influence of the luminous flux incident upon it.

The external-circuit (output) current i_n of a multiplier consisting of n stages may be calculated if we know the primary photocurrent i_0 (at the input) of the multiplier and the secondary emission ratio σ, i.e., the number of secondary electrons dislodged from the emitter by the incidence of one primary electron. Assuming that the secondary

emission ratio is uniform for all emitters, we have

$$i_1 = \sigma i_0;$$
$$i_2 = \sigma i_1 = \sigma(\sigma i_0) = \sigma^2 i_0;$$
$$\cdots\cdots\cdots\cdots$$
$$i_n = \sigma^n i_0,$$

where n is the number of emitters or stages of the multiplier. It follows from this that the multiplier gain is $k = \sigma^n$ and depends on the number of stages n in the multiplier and the secondary electron emission ratio σ.

A simple rough calculation shows that in the presence of good emitters having, for example, a secondary emission ratio $\sigma = 10$, the gain of a ten-stage $(n = 10)$ multiplier is $k = \sigma^n = 10^{10}$, i.e., the primary photocurrent may be multiplied ten billion times with a ten-stage multiplier of this type. In practice, however, it is impossible to achieve such amplifications. The reason for this is primarily the fact that it is impossible to manufacture a multiple-stage multiplier whose secondary emission ratio is sufficiently high and uniform for all emitters. Furthermore, it is impossible to direct (focus) all the electrons emitted by a given intermediate emitter at the following emitter, since some of these electrons pass, for example, from the first to the third emitter, by passing the second, or from the next-to-last to the plate, by passing the last emitter. Finally, some of the electrons are scattered all over the multiplier, striking the wall of the bulb, the supports, etc. The result is a sharp decline in the gain of the multiplier. In practice, it is possible to make multiple-stage multipliers with gains up to $k = 10^8$.

A magnetic or electric field is used to focus electrons in multiple-stage multipliers (magnetic or electrostatic focusing). Magnetic focusing is accomplished by placing the multiplier in the strong mag-

netic field between the poles of a permanent horseshoe magnet. When the lines of force of the magnetic field are directed perpendicular to the velocities of the electrons as they move from one emitter to another, the electrons, accelerated by the electric field and deflected by the magnetic field, will move along cycloidal paths (Fig. 14-16), striking one emitter after another in succession. Thus, the electron scattering is reduced and the gain of the multiplier is increased considerably.

The necessity of using cumbersome magnets or electromagnets is a drawback to magnetic focusing. It is for this reason that photomultipliers with magnetic focusing have not found extensive application.

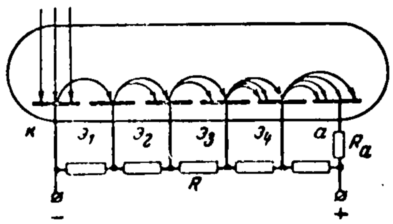

Fig. 14-16. Schematic diagram of multiplier with magnetic focusing.
k — cathode; a — plate; e_1 to e_4 — emitters; R — voltage-divider resistance; R_a — load resistance (the magnetic field is directed perpendicular to the plane of the drawing).

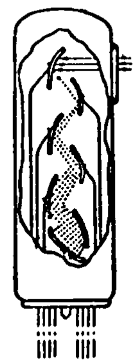

Fig. 14-17. Positions and shapes of emitters in multiple-stage FEU with electrostatic focusing.

Electrostatic focusing is usually achieved by shaping the emitters in such a way as to permit an electric field to form between the emitters and focus the electrons. In this case, the electrons passing from one emitter are focused toward the next emitter by the electric field acting between these emitters. The quality of the focusing effect depends upon the shape of the electric field, which, in turn, is determined by the shape and arrangement of the emitters. Emitters that are approximately semicylindrical in shape, for example, and are positioned

in a definite sequence (Fig. 14-17) have good focusing properties. When the photocathode is of large size, a diaphragm D is placed between it and the first emitter to perform the function of an electronic focusing lens for the stream of primary electrons moving from the photocathode toward the first emitter. This shape and placement of the emitters ensure such good focusing that electrons do not overshoot the emitters, even with large output currents in the last stages. In addition, the detrimental influence of the positive ions created near the plate by collisions of electrons with residual gas molecules is minimized in this multiplier design. In it, positive ions moving toward the photocathode stick to the emitters closest to the plate and, consequently, do not reach the photocathode. In other design types in which the emitters are differently shaped or differently located (for example, do not follow one another), the incidence of positive ions on the photocathode brings about secondary electron emission from it and thereby increases the dark current.

At the present time, there are a large number of different multiple-stage photomultiplier designs with the most diverse parameters and characteristics.

C. The parameters of multiple-stage photomultipliers.

The basic parameters of the multiple-stage photomultiplier are its total sensitivity, gain, dark current, photocathode sensitivity, and plate insulation resistance. The above multiplier parameters are determined by its design and by the properties of the photocathodes and emitters used.

If the total sensitivity of the photocathode is denoted by k_0 and the multiplier gain by k, the total sensitivity K_0 of the multiplier may be given by the equation

$$K_0 = k_0 k. \tag{14-1}$$

The total sensitivity of the photocathodes used in photomultipliers is usually lower than that of those used in photocells. This is due to the fact that the photocathode and emitters of photomultipliers are sensitized in the same mode, while it is advisable to sensitize them separately to obtain high sensitivity, and this is impossible in practice.

The total sensitivities of multiple-stage multipliers are measured in amperes per lumen (or per watt). It does not follow from this, however, that the multiplier's cathode may be illuminated by an intense luminous flux and large currents obtained at the output, since such currents may break down the output emitters. Multiple-stage photomultipliers, on the other hand, are designed for operation at minimal luminous flux. Thus, the sensitivity of the multiplier may reach tens of amperes per lumen, but the greatest attainable output currents do not usually exceed a few ma or, at most, tens of milliamperes. Thus, the function of the multiplier does not consist in using the greatest possible luminous fluxes to produce the greatest possible output currents, but in permitting work with minimal luminous fluxes, at which the output current frequently does not exceed the current of a photocell.

Photocells with extrinsic photoeffects cannot be used at extremely low luminous fluxes (of the order of thousandths or ten thousandths of a lumen) because the photocurrent is too small at such fluxes and effective amplification can be accomplished only by means of secondary electron emission.

At a luminous flux F, the total sensitivity K_0 of the multiplier and the total sensitivity k_0 of the photocathode are related by the equations

$$K_0 = \frac{I_a}{F} \qquad \text{[amp/lumen]};$$

$$k_0 = 10^6 \frac{I_k^*}{F} \qquad [\mu\text{a/lumen}],$$

* $[I_k - I_k - I_{katod} - I_{cathode}.]$

where I_a and I_k denote the currents in the plate and cathode circuits.

The gain k of the multiplier is given by Eq. (14-1):

$$k = \frac{K_0}{k_0}.$$

In contemporary multiple-stage photomultipliers, the gain and total sensitivity reach values of 10^8 and thousands of amp/lumen, respectively, with total photocathode sensitivities of the order of a few tens of microamperes per lumen.

TABLE 14-4.
Basic parameters and typical operating modes of certain types of multiple-stage photomultipliers.

1 Типы умножителей	2 Число каскадов усиления	3 Оптический вход	4 Типы катодов	5 Спектральная область, Å	6 Максимум спектральной характеристики, Å
13 ФЭУ-17	13	19 Боковой	21 Сурьмяно-цезиевый	3 000—6 000	3 900±300
14 ФЭУ-18	13	»	22 То же	2 200—6 000	3 700±300
15 ФЭУ-19М	13	20 Торцовый	» »	3 000—6 000	3 900±300
16 ФЭУ-20	8	19 Боковой	» »	3 500—6 000	3 900±300
17 ФЭУ-22	13	»	23 Кислородно-цезиевый	10 000	7 500±1 000
18 ФЭУ-25	9	20 Торцовый	21 Сурьмяно-цезиевый	3 000—6 000	3 900±300

1	7 Интегральная чувствительность катода, мка/лм	8 Максимальный ток на выходе (постоянный), мка	9 Примерные рабочие режимы: 1 — 10 Напряжение, в	1 — 11 Чувствительность, а/лм	1 — 12 Темновой ток, а	2 — 10 Напряжение, в	2 — 11 Чувствительность, а/лм	2 — 12 Темновой ток, а
13	25	100	775	10	$3 \cdot 10^{-9}$	950	100	$3 \cdot 10^{-8}$
14	25	100	775	10	$3 \cdot 10^{-9}$	950	100	$3 \cdot 10^{-8}$
15	35	200	1 150	10	$3 \cdot 10^{-9}$	1 500	100	$3 \cdot 10^{-8}$
16	35	100	900	2,5	$3 \cdot 10^{-9}$	1 200	10	$1 \cdot 10^{-8}$
17	10	300	1 400	4,5	$1,5 \cdot 10^{-8}$	—	—	—
18	35	100	1 250	4	$6 \cdot 10^{-9}$	1 700	28	$8 \cdot 10^{-7}$

1) Multiplier type; 2) number of amplification stages; 3) optical input; 4) cathode type; 5) spectral region, A; 6) maximum of spectral characteristic, A; 7) total sensitivity of cathode, μa/lumen; 8) maximum output current (constant), μa; 9) typical operating modes; 10) voltage, v; 11) sensitivity, amp/lumen; 12) dark current, amp; 13) FEU-17; 14) FEU-18; 15) FEU-19 M; 16) FEU-20; 17) FEU-22; 18) FEU-25; 19) side; 20) end; 21) antimony-cesium; 22) same; 23) oxygen-cesium.

The so-called sensitivity threshold, i.e., the smallest luminous

flux that is still measurable or detectable by the multiplier (or photocell) in question is an important parameter of multipliers (and photocells). If the luminous flux is reduced still further, the current that it produces becomes of the same order as the dark current and consequently cannot be measured or detected.

The dark current itself, that flowing in the multiplier at the operating voltage but in complete darkness, is no less important a parameter of the multiple-stage photomultiplier.

The dark current consists of the thermionic currents of the photocathode and the first-stage emitters, which are amplified in the subsequent stages, leakage currents across the insulation between the plate and the other electrodes, i.e., across the inside and outside surfaces of the glass bulb and support, and currents due to spontaneous electron emission and to ionic and optical feedback. Positive ions which form in the plate end of the multiplier and move from the plate toward the first emitters and the photocathode in the accelerating field give rise to the ionic feedback. Bombarding the first emitters and the photocathode, these ions bring about secondary emission of electrons, with the result that the dark current also increases. Optical feedback is due to low-grade illumination of the photocathode by the luminescence of the emitters and the glass of the bulb – an effect of the electrons in the plate and of the multiplier. The cathodic emission that arises as a result of the luminescence also increases the dark current. Naturally, the dark current limits the luminous sensitivity threshold of multipliers (and photocells). Thus, for example, the luminous sensitivity threshold of multiple-stage photomultipliers with oxygen-cesium cathodes is $1.4 \cdot 10^{-9}$ lumen, while the luminous sensitivity threshold of photomultipliers with antimony-cesium cathodes is $5.7 \cdot 10^{-12}$ lumen.

The basic parameters of certain types of multiple-stage photomulti-

pliers appear in Table 14-4.

D. Characteristics of multiple-stage photomultipliers.

The basic characteristics of a multiple-stage photomultiplier are the gain k as a function of the total voltage U of the multiplier's power source, i.e., k = f(U) and the light- and volt-ampere characteristics.

If the gain k is plotted against the Y axis on a logarithmic scale, the functional relationship k = f(U) may be represented by curve a for a multiplier with electrostatic focusing and by curve b for a multiplier with magnetic focusing (Fig. 14-18). These curves show that as the source voltage is increased from 250 to 1250 v, the gain of the electrostatically-focused multiplier increases continuously, while the gain of the magnetically-focused multiplier rises to a maximum and then drops off somewhat as the source voltage increases further. It is apparent from this that there is no necessity for strong electric fields with magnetic focusing, since it appears that the fastest electrons emitted by the next-to-last emitter proceed directly to the plate under the influence of the strong electric field, bypassing the last emitter so that the gain drops off slightly.

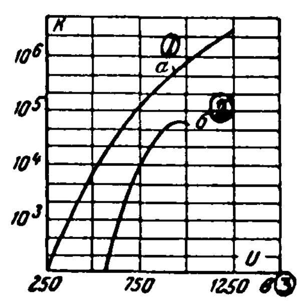

Fig. 14-18. Curves of gain as a function of source voltage for multiple-stage FEU. 1) For electrostatic focusing; 2) for magnetic focusing; 3) volts.

The light characteristic of a multiple-stage multiplier is represented by the curve of the plate current I_a as a function of the lumin-

ous flux F striking the photocathode (Fig. 14-19). As is shown by the diagram, the light characteristic is linear for the segment from 10^{-12} to 10^{-3} lumen. The linearity breaks down at higher luminous fluxes. This is accounted for by the fact that the multiplier gain drops off at high luminous fluxes as a result of emitter fatigue and because of the influence exerted by the space charges that form in the last stages and the plate end of the multiplier.

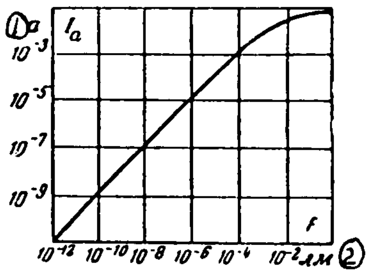

Fig. 14-19. Light characteristic of multiple-stage FEU.
1) amp; 2) lumens.

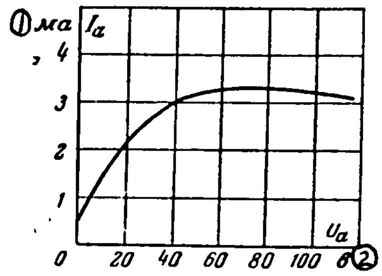

Fig. 14-20. Plage characteristic of multiple-stage FEU.
1) ma; 2) volts.

The plate characteristic of a multiple-stage multiplier, which expresses its plate current I_a as a function of the voltage U_a applied between the plate and the last emitter with the steady-state voltage distribution among the rest of the emitters, is given in Fig. 14-20. This characteristic shows that when saturation has been reached, a small plate-current drop takes place for high plate voltages due to the incidence of a certain fraction of the electrons from the next-to-last emitter directly onto the plate without striking the last emitter.

The basic parameters of certain types of multiple-stage photomultipliers are shown in Table 14-4.

14-6. APPLICATIONS OF PHOTOCELLS AND PHOTOMULTIPLIERS.

By virtue of the fact that they have practically instantaneous response, vacuum-tube photocells with extrinsic photoeffect and photo-

multipliers have found extensive and diversified applications in the most diverse fields of science and engineering. They are capable of operation both on modulated luminous fluxes (sound-on-film, sound-on-paper, facsimile telegraph, etc.), and on luminous fluxes that vary discontinuously or qualitatively (in automatic-control systems).

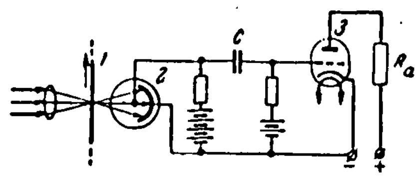

Fig. 14-21. Schematic diagram of reproduction of sound recorded on movie film. 1) Film with recording; 2) photocell; 3) amplifier input tube.

Photocells with extrinsic photoeffect and single-stage photomultipliers are extensively used in sound-film work for reproduction of sound from light traces. A schematic diagram of sound reproduction is given as an example in Fig. 14-21. The light focused by the optical system shines through the moving film in regions where sound has been recorded. Having passed through the film, the light strikes the cathode of the photocell, giving rise to the corresponding emission of photoelectrons. As a result, a current arises in the external circuit and varies as a result of the variation of the luminous flux incident upon the photocathode in accordance with the varying photographic density (transparency) of the sound track recorded on the film. For this reason, a varying voltage drop arises across the load resistance R_a of the external photocell circuit and is applied through the capacitor C to the grid of the amplifier input tube, while the plate current at the amplifier output varies in accordance with the variation of the voltage across the load resistance. It is this current that feeds the loudspeaker.

Many types of photocells are extensively used in the various types of photorelays, in which very small photocurrents are used to control the relatively large currents that power various electrical or mechanical devices. For example, Fig. 14-22 shows a schematic diagram of a relay with a photocell and a gas-discharge tube. A potential insufficient to fire the tube is applied at the point where the tube and the photocathode are connected, and when the photocell is illuminated, the potential of the common point declines and the potential difference between the electrodes of the gas-discharge tube increases, so that the tube ignites and the relay operates.

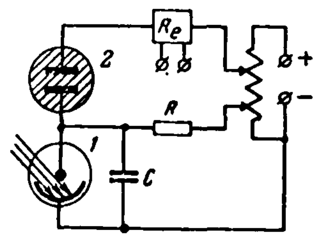

Fig. 14-22. Schematic diagram of photorelay.
1) Photocell; 2) gas-discharge tube.
R_e — relay; R — resistance;
C — capacitance.

The current at the relay output is usually used for the purpose of switching lights on and off, starting and stopping motors, or to drive counters to count the number of objects passing between the photocell and the light source. It is possible to use photorelays to switch electric lighting systems on and off automatically in cities and large industrial centers, adjusting them to operate when the intensity of the natural illumination rises or falls to a definite value.

Multiple-stage photomultipliers have recently come into extensive use in science and engineering. They are used in astronomy for photometry of the solar and stellar spectra and in atmospheric optics for

study of the infrared radiation of the nighttime sky, the aurora, and other objectives. Multiple-stage photomultipliers are also being used successfully in spectral analysis (for example, in studies of the spectral and time characteristics of luminophors). Special designs in which photomultipliers are used in combination with luminescent screens have been used on a particularly broad scale to detect individual charged particles by the light flashes (scintillations) that occur when luminescent objects are bombarded by the particles under investigation.

Objective monitoring of light sources and the temperatures of heated bodies, spectral analysis of materials, and monitoring of the intensity variations of natural or artificial illumination are functions that can be performed with the aid of photocells and photomultipliers. Photocells and photomultipliers are also being used successfully as color analyzers, where they permit accomplishment of color control, grading of colored materials, and control of signal-light systems. They are also used in safety devices that operate on invisible (for example, infrared) as well as visible light, thus permitting monitoring functions to be carried out in darkness.

Chapter Fifteen

CATHODE-RAY AND ELECTROOPTICAL DEVICES

15-1. CLASSIFICATION OF CATHODE-RAY AND ELECTROOPTICAL DEVICES.

The class of cathode-ray devices includes devices utilizing the energy of a focused pencil of electrons, normally called an electron beam.

Below we shall consider devices utilizing a focused electron beam that may be classified into the following groups:

1. Cathode-ray tubes that convert electric signals into a visible image. This group of devices includes oscilloscope tubes designed for studying alternating electrical processes, television picture tubes, and tubes employed for the screens of radar installations.

2. Transmitting television tubes, designed to convert a visible image into electric signals.

3. Electrooptical converters used to convert a light image into an electronic and then back to a light image again, but with a different spectral composition.

Electron microscopes and electron-beam switches operating with focused electron beams are not considered in this book.

15-2. ELECTRON-BEAM TUBE STRUCTURE. ELECTRON GUNS.

A cathode-ray tube consists of the following sections:

1) an electron gun, which creates a narrow (with small cross-sectional area) electron beam of the required intensity, which may be varied within predetermined limits;

2) a deflection system the function of which is to change the direction of the electron beam;

3) a screen, glowing or changing its color (absorption spectrum) under the action of the beam of electrons. In the majority of cases,

the screen is applied to the inside surface of the envelope of the tube, located opposite the electron gun and is a layer of a phosphor — a material that glows under the action of electron bombardment. This

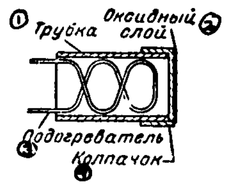

Fig. 15-1. Cathode structure for electron-beam tube.
1) Tube; 2) oxide layer;
3) heater; 4) cap.

glow is observed through the glass of the envelope.

The dimensions of the light spot on the screen are smaller the smaller the area of the beam cross-section, i.e., the better it is focused. The brightness with which the screen glows at the point of incidence of the beam depends upon the magnitude of the current, and upon the energy of the electrons. By changing the direction of the beam in the device with the aid of the deflection system, it is possible to shift the bright spot on the screen, while by changing the magnitube of the beam current with the energy of the electrons the brightness

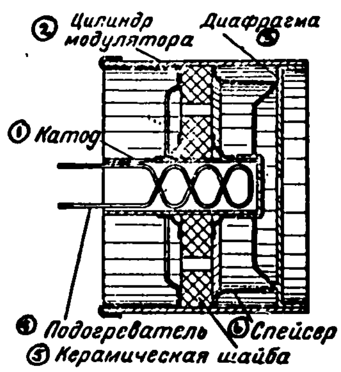

Fig. 15-2. Typical cathode-modulator unit, in section.
1) Cathode; 2) modulator cylinder; 3) diaphragm;
4) heater; 5) ceramic bead;
6) spacer.

may be varied.

For a given electron energy, the construction of the electron gun should provide the required size of bright spot and magnitude of beam current. In addition, the characteristic of an electron gun is a modulation characteristic determining the nature of the variation of the beam current within the required limits.

An electron gun consists of a hot cathode, a control electrode (modulator), and a system of electrodes that focus the electron beam on the screen. In the majority of cathode-ray tube designs, an indirectly heated oxide-coated cathode is used; it takes the form of a tube with a cap on the outside surface of the bottom of which there is applied an oxide layer (Fig. 15-1). In the majority of designs in use, the cathode is located within a cylinder that has a hole in its bottom (in the center), forming a diaphragm. This cylinder serves as the modulator; by changing its potential with respect to the cathode, it is possible to vary the number of electrons passing through the diaphragm, i.e., it is possible to regulate the magnitude of the beam current. Sometimes, the modulator cylinder projects beyond the plane of the diaphragm, forming a so-called modulator "skirt" (Fig. 15-2).

The electrons coming off at different points on the cathode have their initial velocities directed variously. Under the action of the electric field set up at the surface of the cathode by the potentials on the modulator (normally negative) and the accelerating electrode located beyond the diaphragm, these electrons pass through the aperture in the diaphragm, or are returned to the cathode. The maximum field strength of the accelerating field obtains in the region opposite the center of the diaphragm aperture, and it decreases in the radial direction, and as a result, the maximum current density is taken from the central section of the cathode. When the modulator potential is

increased, the cathode current rises, first, owing to an increase in the cathode surface yielding electrons and, second, owing to the increased cathode-current density. Figure 15-3 gives sample curves for the distribution of current density from a cathode for three values of

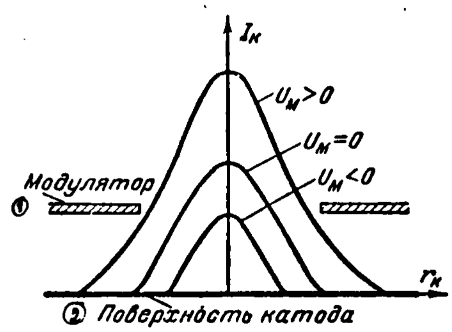

Fig. 15-3. Distribution of current density on a cathode surface with various modulator potentials.
1) Modulator; 2) cathode surface.

modulator potential.

In practice, the maximum current density taken from the central region of the cathode amounts to 0.2-0.5 amp/cm^2; in certain types of devices, it reaches 1 amp/cm^2.

In choosing the dimensions of the cathode-modulator unit of a gun, it is essential to know the way in which the magnitude of the cathode current depends upon the modulator potential and the accelerating-electrode potential. To determine the cathode-current magnitude, it is possible to make use of the following empirical formula:

$$I_к = \frac{3(U_M - U_{M0})^{3.5}}{U_{M0}^2}, \qquad (15\text{-}1)$$

where I_k is the cathode current, μa; U_M is the modulator potential, v; U_{MO} is the magnitude of modulator potential at which the beam current is zero; this value is called, as in the case of an electron tube, the cut-off voltage.

The magnitude of the cut-off voltage depends upon the geometry of the electrodes and the potential of the accelerating electrode. It is considerably more complicated, however, to calculate the magnitude of the cut-off voltage in an electron gun than it is for an electron

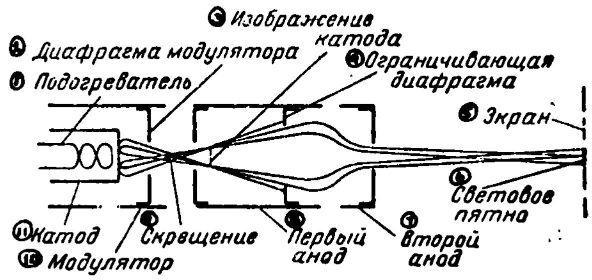

Fig. 15-4. Very simple electron gun with electrostatic beam focusing.
1) Heater; 2) modulator diaphragm; 3) image of cathode; 4) limiting diaphragm; 5) screen; 6) light spot; 7) second anode; 8) first anode; 9) cross-over; 10) modulator; 11) cathode.

tube. Experimental data for systems in practical use as cathode-ray-tube guns give the following approximate expression connecting the magnitude of the cut-off voltage with electrode dimensions and the accelerating-electrode voltage:

$$U_{M0} = k\frac{D^3}{\delta r_{Mk}\, r_{Ma}}\, U_{a1}. \tag{15-2}$$

In this expression, D is the diameter of the aperture in the modulator diaphragm; δ is the diaphragm thickness; r_{Mk} is the cathode-modulator separation; r_{Ma} is the spacing between the modulator diaphragm and the accelerating electrode; U_{a1} is the accelerating-electrode voltage. The proportionality constant k depends upon the dimensions of the modulator skirt and the design of the accelerating electrodes.

The stream of electrons passing through the modulator diaphragm

must be focused on the screen of the tube. In order to do this, the electron gun has a focusing system normally consisting of two electron lenses.

Where electrostatic focusing of the electron beam is used, the focusing system normally consists of two electrostatic lenses, made in the form of diaphragms (Fig. 15-4), or in the form of coaxial cylinders of the same or different diameters.

The first lens of a focusing system is formed by three electrodes: a cathode, a modulator, and a first anode. The electrons, leaving the cathode with various initial-velocity directions undergo axial and radial acceleration; in this case, the radial acceleration is directed towards the axis. Since the initial velocities of the electrons are low, the electron trajectories are curved sharply toward the axis of the tube, and the electron beam is compressed near the modulator, after which it again diverges. The region corresponding to minimum electron-beam diameter is called the intersection. In the intersection region, the trajectories of the electrons cross if they come from different points on the cathode, while the trajectories of electrons coming from the same cathode point intersect considerably further out, in a plane, forming a plane image for the first length. The size of the cathode image, as is shown on the figure, is larger than the width of the beam in the intersection region. Thus, in order to decrease the size of the light spot, the second lens is so arranged that the image of the intersection, rather than the image of the cathode, is projected on the screen. Consequently, electron paths should intersect on the screen when the electrons have passed through a single point of the intersection.

The second electrostatic lens, called the final lens, is formed by the field between the first and second anodes. Since in the final

lens field, the electrons arrive at high velocities, in order to focus them, a rather high potential gradient is needed. Thus, the voltage on the second anode should be considerably higher than the voltage on the first anode. As a rule, the first-anode voltage is about 10-30 per cent of the second-anode voltage. In addition, in the majority of cathode-ray devices, with the exception of tubes with after-acceleration, the magnitude of the second-anode voltage determines the energy with which the electrons arrive at the screen.

The anodes have diaphragms which limit the angle of divergence of the electron beam, cutting out from the entire beam those electrons that are moving at a relatively small angle with respect to the axis. This improves focusing of the electron beam, and decreases the size of the light spot. Owing to this clipping of some of the electrons by the anode diaphragms, the beam current arriving at the screen turns out to be less than the cathode current.

Focusing of an electron beam is achieved by varying the voltage on the first anode. In other words, by changing the voltage on the first anode, it is possible to obtain a field in the lenses that will produce the least light-spot area.

More common are systems of electron guns using an accelerating electrode. In such an electron-gun system, there is an additional accelerating electrode, located between the modulator and the first anode, which is at a high positive potential, as a rule higher than the first-anode potential. In tubes with electrostatic focusing (with the exception of high-voltage tubes), the accelerating electrode is frequently connected directly to the second anode (Fig. 15-5). In the majority of modern tubes with electrostatic beam focusing, electron guns are used that have zero first-anode current, which makes it possible to supply the anode from a common source through a voltage divider.

The peculiarity of such guns is the form of the accelerating electrode and second anode, which are cylinders with limiting diaphragms, while the first anode is a diaphragm or cylinder having an aperture of large diameter. In this gun, the cathode, modulator, and accelerating elec-

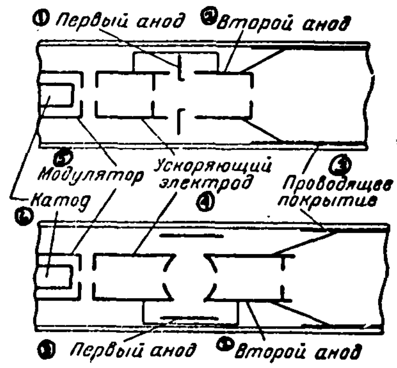

Fig. 15-5. Variant designs for electron guns with accelerating electrodes.
1) First anode; 2) second anode; 3) conducting coating; 4) accelerating electrode; 5) modulator; 6) cathode.

trode form the first, immersion, lens of the gun, while the system consisting of the accelerating electrode and the second anode, electrically interconnected, form, with the aid of the first anode between them, the second separate lens.

Figure 15-6 shows one of the versions of such an electron gun. In assembling the gun, it is especially important to ensure that the electrode axes coincide precisely, and that the spacings between them are maintained, with primary attention to the cathode-modulator spacing. In order to provide accurate location of the cathode and modulator with respect to each other, the cathode is embedded in a ceramic bead, tightly inserted in the modulator cylinder. The distance between the oxide-coated cathode surface and the modulator diaphragm, fixed with the aid of a spacer, is on the order of 0.12-0.16 mm. A decrease in

the spacing is limited by possible thermal deformations of the cathode and modulator diaphragm leading to a large relative change in this distance, extending even to a short circuit. The diameter of the diaphragm aperture in the modulator ranges from 0.5 to 1.0 mm, depend-

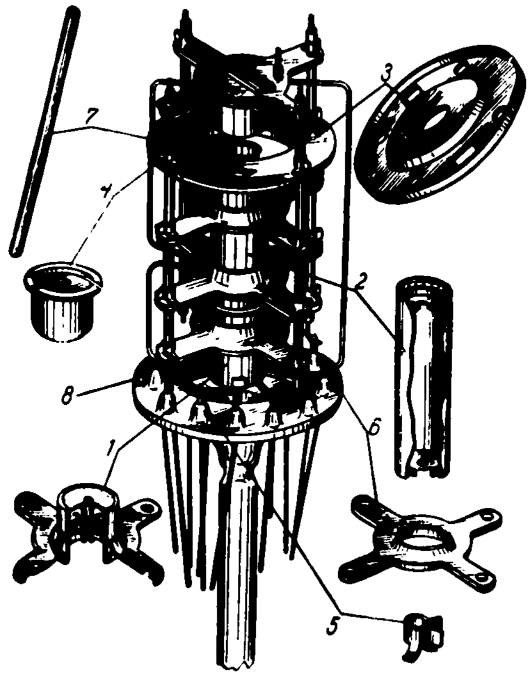

Fig. 15-6. Structure of electron gun.
1) Cathode-modulator unit; 2) accelerating electrode; 3) first anode; 4) second anode; 5) supports; 6) spider for mounting electrode; 7) ceramic tube; 8) stem.

ing upon the value of the accelerating-electrode voltage, and the required magnitude of beam current and cut-off voltage. We ensure that the electrodes of the gun have the same axis by assembling them on a special chuck. In this case, the spacing is fixed by means of pads that are removed after assembly. The entire system is fastened to ceramic beads with the aid of either vacuum cement, or metal supports.

Any coaxial deviation of the parts will disturb the coaxiality of the electric fields that form the lenses, and, consequently, will distort electron-beam focusing.

61-102

Even where all of the electrodes making up the electron gun are accurately coaxial, however, in practice, it is not possible to obtain a sharp intersection image on the screen. Image errors, called aberra-

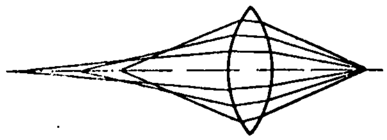

Fig. 15-7. Representation of spherical aberration, as exemplified by a glass lens.

tions, are caused by many factors connected with imperfection of the optical system. One of the most substantial sources of image distortion is so-called spherical aberration, which occurs both in electro-optical and in glass lenses. Beams (or electrons), coming from a single point on the focused subject (or in the case of an electron gun, from a single point of the cathode) at different angles, do not converge to a single point; as a result, a diffused image of the subject is obtained, as shown for the case of a glass lens in Fig. 15-7. It is clear that the magnitude of spherical aberration will be less the smaller

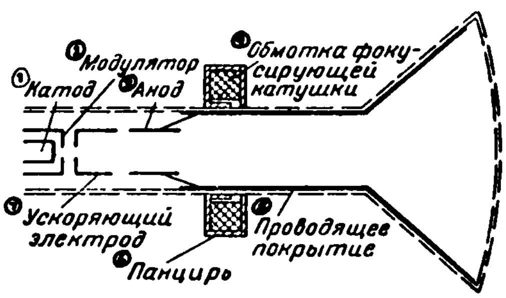

Fig. 15-8. Electron gun with mixed focusing.
1) Cathode; 2) modulator; 3) anode; 4) focusing-coil winding; 5) conducting coating; 6) jacket; 7) accelerating electrode.

the dispersion angle of the beam of electrons. Thus, in order to de-

crease spherical aberration, limiting (aperture) diaphragms are used that cut off a section with small dispersion angles from the entire electron beam.

In many types of cathode-ray tubes, electron guns are used that have mixed electrostatic and magnetic focusing (Fig. 15-8). In such guns, the first lens is electrostatic and the second magnetic. The advantage of the magnetic lens in comparison with the electrostatic lens lies in the smaller value of spherical aberration, which results from the small diameter of the lens in comparison with an electrostatic lens.

As in a gun with electrostatic focusing, the first lens is made with an additional accelerating electrode. A positive potential on the order of 200-400 v DC is applied to the accelerating electrode. The anode voltage reaches several thousand volts.

The second, magnetic, lens is normally set up by a magnetic coil placed on the cylindrical neck of the tube envelope, and encased in an iron jacket. Such coils form a lens with a relatively short focus. In installing a focusing coil, it is essential that it be accurately centered with respect to the optical axis of the tube.

15-3. DEFLECTING SYSTEM OF CATHODE-RAY TUBES.

The electron beam, focused by the gun, is directed toward the center of the screen of the tube. In order to force the light spot to move across the screen, the tubes have deflecting systems that change the direction of the focused electron beam, acting on it by means of a transverse electrical magnetic field.

In the simplest case, electrostatic deviation is accomplished by a system that consists of two pairs of parallel plates, located so as to be mutually perpendicular (Fig. 15-9). If there is a potential

difference between any pair of plates, the electrons of the beam will be deflected toward the plate with the higher potential. The paths of motion of the electrons in the field of the plate, which may be con-

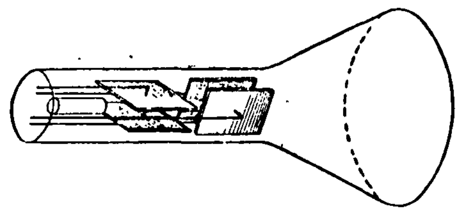

Fig. 15-9. Location of deflecting plates.

sidered to be uniform in first approximation, are parabolic. Passing through the field of the plates, the electrons move along a straight line tangent to the parabola at the point where the electrons exit from the field (Fig. 15-10). As a result, the beam strikes the screen not in its center, but at some distance h from the center. The quantity h, called the deflection, may be found from the known geometric dimensions of the tube elements, the velocity of the electrons, and the potential difference between the deflecting plates.

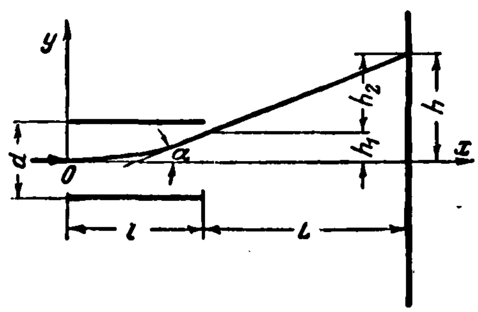

Fig. 15-10. Derivation of deflection sensitivity.

Let us consider the motion of an electron moving in the field of the deflecting plates at a velocity directed along the optical axis. If a potential difference U is applied to the plates, the field strength between the plates equals:

$$E=\frac{U}{d}. \tag{15-3}$$

Choosing as the origin of the coordinate system the point at which an electron enters the field of the plates, and directing the X axis along the optical axis of the tube, and the Y axis perpendicular to the plates, it is possible, in accordance with Eq. (2-7), to write the equation of motion for electrons in the field of the plate:

$$y=\frac{eE}{2mv_0^2}x^2=\frac{eU}{2mdv_0^2}x^2. \tag{15-4}$$

The velocity of an electron v_0, acquired by it in the electron gun, equals:

$$v_0=\sqrt{2\frac{e}{m}U_a},$$

where U_a is the voltage on the anode of the gun.

Substituting this value v_0 into Eq. (15-4), we obtain:

$$y=\frac{U}{4dU_a}x^2. \tag{15-4a}$$

The true magnitude of the deflection h equals the sum of the deflection of an electron in the field of the plates h_1 and the deflection h_2 caused by the change in direction of motion of an electron in the field of the plates. h_1 equals the coordinate y of the electron at the point where it leaves the plate field, i.e., at x = l (length of the plate):

$$h_1=\frac{U}{4dU_a}l^2. \tag{15-5}$$

The deflection h_2 depends upon the separation of the plate from the beam and the angle of inclination of the tangent to the parabola at the point where the electron leaves the field of the plates.

$$h_2=L\,\mathrm{tg}\,\alpha. \tag{15-6}$$

The magnitude of tan α is determined from Eq. (2-10)

$$\operatorname{tg}\alpha = \frac{eEl}{mv_0^2}.$$

Substituting this value of tan α into (15-6), and expressing E in terms of U, and v_0 in terms of U_a, we obtain an expression defining h_2:

$$h_2 = L\frac{Ul}{2dU_a}. \tag{15-7}$$

The total value of the deviation h equals:

$$h = h_1 + h_2 = \frac{U}{4dU_a}l^2 + L\frac{Ul}{2dU_a}, \tag{15-8}$$

or

$$h = \frac{Ul}{2dU_a}\left(\frac{l}{2} + L\right). \tag{15-8a}$$

Letting $l/2 + L = D$ (D is the distance from the screen to the center of the plates), we obtain, finally:

$$h = \frac{lD}{2dU_a}U. \tag{15-9}$$

It follows from expression (15-9) that the magnitude of the beam deflection is directly proportional to the voltage applied to the deflecting plates. The ratio

$$\frac{h}{U} = \frac{lD}{2dU_a} = k(\mathrm{mm/v}) \tag{15-10}$$

is called the deflection sensitivity. The magnitude of the sensitivity is defined as the deflection of the light spot per volt of deflection voltage. Expression (15-10) shows that, first, the sensitivity is inversely proportional to the plate voltage, which determines the velocity of the electrons and, second, that the sensitivity is independent of the electron mass and charge and, consequently, the deviation will be the same for any charged particles leaving the gun.

The expression for the deviation of the light spot was derived with the condition that the electric field of the deflection plates is uniform. The plate field, however, may be assumed to be uniform only at the center of the plate region (Fig. 15-11). Owing to the edge effect,

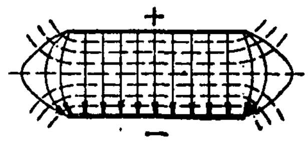

Fig. 15-11. Electric field in the space between deflecting plates (lines of force are shown solid, and equipotential lines by the dashed lines).

the deflecting field extends beyond the plates themselves, which somewhat increases the sensitivity. For the plate dimensions used in practice, this increase in sensitivity is on the order of 15 per cent.

If an alternating voltage is applied to the deflecting plates, the position of the light spot on the screen changes continuously with a velocity corresponding to the rate of change of the deflecting voltage; the spot will shift in a direction perpendicular to the planes of the plates. Since in the tubes there are two mutually perpendicular pairs of deflecting plates, the light spot will be shifted in two

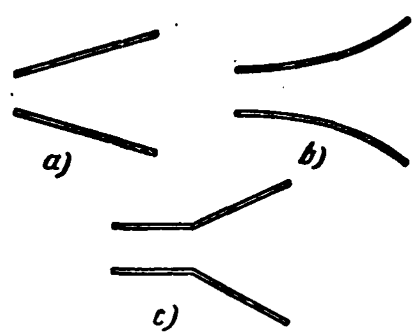

Fig. 15-12. Deflecting plates. a) Nonparallel plates; b) bent plates; c) parallel plates with linearly diverging ends.

mutually perpendicular directions and, consequently, may occupy any position on the screen, determined at each moment of time by the magnitudes of the deflecting voltages.

In many cathode-ray tubes, it is necessary to obtain a large deflection of the light spots, which corresponds to a large angle between the extreme positions of the electron beam, which is called the deflection angle. A large deflection angle makes it possible to decrease the difference from the screen to the deflecting plates without changing the size of the screen and, consequently, decreases the dimensions of tubes and equipment in which they are used. For plane-parallel deflecting plate, however, an increase in deflection angle while maintaining sensitivity is limited by the cutting of the electron beam by the plates themselves.

Thus, plane-parallel plates are utilized in cathode-ray tubes rather seldom. Plane plates that are not parallel are used more frequently, as are parallel plates with linearly diverging edges, or sometimes curved plates (Fig. 15-12).

Both pairs of deflecting plates are located in series, and, consequently, at different distances from the screen. Clearly, if all other conditions are equal, the sensitivity will be greater for the pair of plates that is further from the screen. Thus, for the deflection sensitivity to be approximately the same for both deflection directions, the plates that are further from the screen as a rule are made shorter.

The most important requirement for a deflecting system is the preservation of beam focusing under considerable deflection. This requirement, however, is not always met, and as a result, the size of the light spot will be greater when it is shifted to the edges of the screen than when it is in the center. Even where the field between the deflection plates is ideally uniform, the electron beam is defocused

as a result of its deflection (Fig. 15-13), and the beam defocusing is the greater the wider the beam where it leaves the field of the plates.

Still greater defocusing is caused by the marked nonuniformity

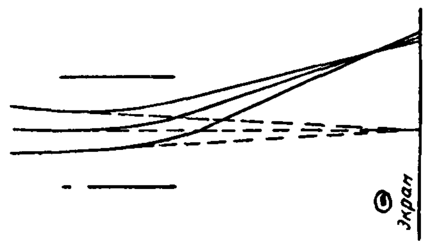

Fig. 15-13. Defocusing of a beam in deflection. 1) Screen.

of the electric field at the edges of the plates. The nonuniform field at the plate edges takes the form of a peculiar electrostatic lens, not having axial symmetry; as a result, electrons moving at various angles to the tube axis are not identically deflected.

Consequently, in order to decrease beam defocusing, it is necessary, first, to decrease beam width at the entrance to the deflecting plates and, second, to decrease field nonuniformity at the edges of the plates.

A decrease in electron-beam width is achieved by installing aper-

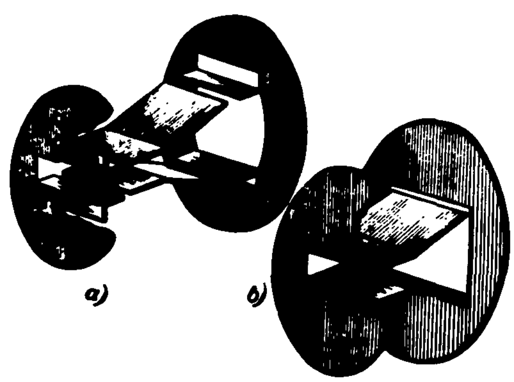

Fig. 15-14. Methods of decreasing field nonuniformity at the edges of deflecting plates.

61-102

ture diaphragms in the output region of the gun anode, which cut out the extreme electrons from the beam. This method, however, decreases the beam current and, consequently, decreases screen brightness as well.

The most common method for supplying the deflecting plates is the so-called symmetric-supply method, in which both deflecting plates of a pair are connected to the gun anode by equal high resistances (on the order of several megohms). When the deflection voltage is applied to the plates, the potential of one of the plates (the positive plate) is higher and that of the other lower than the potential of the anode by an amount equal to one-half of the deflection voltage. The potential of the mean equipotential plane always remains equal to the anode potential. The field at the edges of the plates when this method of supply is used turns out to be symmetric with respect to a plane passing through the axis of the tube, which somewhat decreases deflection distortion.

In addition, in order to decrease field nonuniformity at the edges of the plates, special plates are located near the deflecting plates and they are connected directly to the anode of the gun (Fig. 15-14a). A similar effect may be obtained by placing an electrode near the edges of the deflecting plate that is at the anode potential, and that has a slot with a width equal to the separation between the plates (Fig. 15-14b).

An electron beam may also be deflected by a transverse magnetic field set up by special deflecting coils located outside the tube envelope. Let us consider the deflection of a beam of electrons by a uniform lateral magnetic field. Figure 15-15 shows the deflection of an axial beam in a uniform magnetic field (the field strength vector is directed perpendicular to the plane of the drawing).

As is known, an electron falling into a uniform magnetic field at a velocity U_0 that is perpendicular to the field strength moves in a circle whose radius is determined from the expression (4-20) [sic]:

$$r = \frac{mv_0}{eH}.$$

Consequently, an electron beam in the field of a deflection coil moves along the arc of a circle whose radius is the smaller the greater the magnetic field strength. After leaving the field of the coils, the beam moves toward the screen along the tangent to the arc at the point of exit, and strikes the screen at a point distant from the center. Clearly, the magnitude of the deflection of the light spot from the center of the screen is greater the smaller the radius of curvature of the electrons in the deflection field, i.e., the greater the magnetic field strength. For small deflection angles, the magnitude of deflection is related to the field strength of the deflecting field by the following equation:

$$h = \frac{\sqrt{\frac{e}{2m}}\, aD}{\sqrt{U_a}} H, \tag{15-11}$$

i.e., the magnitude of deflection of the light spot from the center of the screen is directly proportional to the value of the deflecting field strength. At large deflection angles, the relationship between the deflection and the field strength becomes nonlinear.

It is clear from expression (15-11) that the magnitude of beam deflection in a uniform magnetic field, in contrast to electrostatic deflections, depends upon the ratio of charge to mass of the deflected particle. In addition, with magnetic deflection, the magnitude of deflection is inversely proportional to the square root of the accelerating voltage U_a, while for electrostatic deflection, the deflection is inversely proportional to the first power of U_a. This means that with

magnetic deflection, the deflection sensitivity is less dependent upon the accelerating voltage than is the case with electrostatic deflection.

The deflection coils that set up the magnetic field are located on the neck of the tube envelope (Fig. 15-16). As the figure shows, the

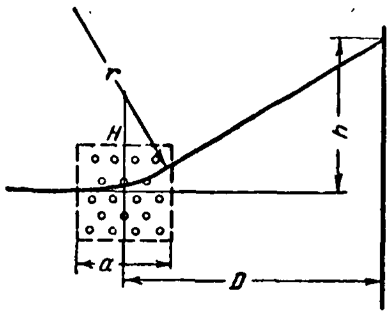

Fig. 15-15. Deflection of electron beam in magnetic field.

coils are located at the same distance from the gun, and the magnetic fields set up by them are superimposed. The mutual superposition of two mutually perpendicular magnetic fields, however, does not cause distortion in beam focusing or deflection. Each pair of deflecting coils sets up a magnetic flux with a field strength directed perpendicular to the tube axis. For a large concentrated magnetic flux, cores and pole pieces of steel are sometimes utilized in the coils. As the frequency of the alternating deflection voltage increases, however, the losses in the steel rise sharply in such coils. Thus, in

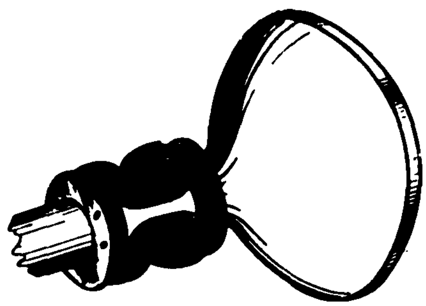
Fig. 15-16. Deflecting coils on neck of cathode-ray tube.

61-10v

many cases, deflection coils are used that do not have cores; their advantage lies in lower weight and the fact that it is more easily possible to obtain the required shape of field by proper choice of the distribution of coil turns.

As with electrostatic deflection, magnetic deflection may cause distortion of the beam focusing and of deflection itself owing to nonuniformity of the magnetic field at the edges of the deflecting coils and the very wide beam. Equalization of magnetic-field uniformity across the tube diameter is achieved by a special distribution of the turns of the deflection-coil windings in which the turns density follows approximately a cosine law from side edges to center. In order to decrease field nonuniformity at the end edges of the coils, their ends are normally bent up.

Owing to the large diameter of the deflecting coils (with respect to the dimension of the beam), the effect of field nonuniformity at the edges of the deflection system turns out to be quite a bit less for magnetic deflection than with electrostatic deflection which is a substantial advantage of the magnetic deflection system. A very important positive quality of magnetic deflection is the fact that sensitivity is less dependent upon accelerating voltage than is the case with electrostatic deflection. In addition, tubes with magnetic deflection are simpler to manufacture and are cheaper. A drawback to magnetic deflection is the fact that the amount of deflection depends upon the mass and charge of the particles deflected, and its effect becomes manifest where there are negative ions in the beam of the tube; we give a more detailed discussion of this case below. The drawbacks to magnetic deflection systems also include the cumbersomeness of the deflection arrangement, as well as the limitation placed upon the permissible frequency of the deviating voltage by coil inductance. In practice, the

frequency of the deflection voltage cannot exceed 10 Mc with magnetic deflection.

15-4. ELECTRON-BEAM SCANNING.

If an alternating voltage is applied to the deflecting plates of a cathode-ray tube, the amount of deflection of the bright spot will vary continuously, i.e., the bright spot will be shifted on the screen in the direction of the deflection. The rate of translation of the spot depends upon the rate of change of the deflecting voltage (or with magnetic deflection, upon the rate of change of the current in the deflecting coils). If the deflecting voltages vary slowly, an observer can differentiate individual positions of the moving bright spot on the screen. When the deflecting voltages change rapidly, the observer, owing to the persistence of human vision which retains a visual impression for a period of about 0.1 sec, will see the motion of the bright spot over the screen as a bright line.

The position of the bright spot on the screen at any moment of time is a function of the instantaneous values of the deflecting voltages, while the form of the bright line depends upon the nature of the change in these voltages. By a suitable choice of the character and magnitudes of the deflecting voltages, it is possible to obtain upon the screen any shape of the bright line. The production of bright lines of required shape is known as electron-beam scanning.

The shape of such a sweep depends upon the function of the cathode-ray tube. Basically, a sweep is produced by applying to the deflecting systems voltages or currents changing according to a predetermined character. In certain tubes, specially designed deflection systems are utilized that produce the required electron-beam scanning. Let us consider certain frequently utilized types of sweeps.

a) Linear sweep. In order to obtain this type of scanning, a periodic sawtooth voltage is applied to one pair of deflecting plates (or coils) (Fig. 15-17); this type of signal gives a linear variation

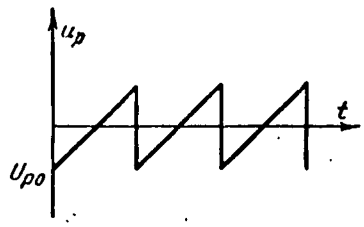

Fig. 15-17. Sawtooth sweep voltage.

in voltage over a cycle, with a nearly instantaneous fly-back to the initial value at the end of the cycle. Under the action of this sweep voltage, the beam is swept across the screen in straight-line segments whose ends are determined by the values of voltage at the beginning and end of a cycle. For a single cycle of sawtooth voltage, it is possible to express these voltages as linear time functions:

$$u_p = at + U_{p0}, \tag{15-12}$$

where a is the rate of change of the sweep voltage, v/sec; U_{p0} is its value at the initial instant ($t = 0$).

If an alternating voltage $u = \varphi(t)$ is applied to the second pair of deflecting plates, it will shift the bright spot in a direction perpendicular to the straight line obtained. Let us use a rectangular coordinate system with the origin in the center of the screen. The X axis will be directed along the line generated by sweeping the beam with the sawtooth voltage. Then the coordinates of the bright spot at any instant of time t (within a single cycle of sweep voltage) will have the following values:

$$x = k_1 u_p = k_1 (at + U_{p0}), \tag{15-13}$$

$$y = k_2 u = k_2 \varphi(t), \qquad (15\text{-}14)$$

where k_1 and k_2 are the sensitivities of both pairs of plates. Solving Eq. 15-13 for t and substituting into Eq. 15-14, we obtain the equation

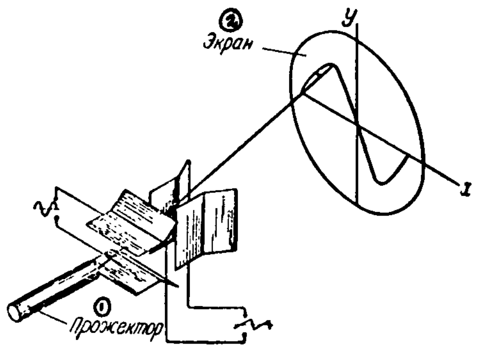

Fig. 15-18. Linear sweep with sinusoidal voltage.
1) Electron gun; 2) screen.

of motion of the bright spot over the screen:

$$y = k_2 \varphi\left(\frac{1}{k_1 a} x - \frac{U_{p0}}{a}\right). \qquad (15\text{-}15)$$

It follows from Eq. 15-15 that the path of motion of the bright spot over the screen will duplicate the shape of the relationship between the voltage u and time.

If the voltage $u = \varphi(t)$ it is also periodic, if the periods of both voltages are equal, the bright spot will describe a curve that corresponds to a single cycle, and that will return to the initial point, and repeat its motion along this curve. In this case, there will be a stable image on the screen of a single cycle of the voltage applied to the second pair of plates. Figure 15-18, such scanning of a sinusoidal voltage. If the frequency of the sawtooth scanning voltage u_p is twice the period of the voltage u, an image of two cycles will appear on the screen, while if the frequency of the first voltage is three times that of the second, there will be three cycles on the screen, etc.

The period of the sawtooth voltage must be equal to or an integral number of times that of the voltage applied to the second pair of plates. If this is not the case, it will be impossible to obtain a stable representation of the second, studied, voltage on the screen.

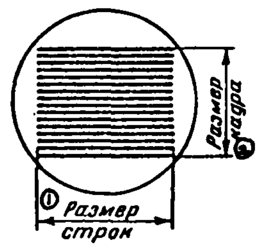

Fig. 15-19. Electron beam scanning of a straight-line raster. 1) Frame width; 2) frame height.

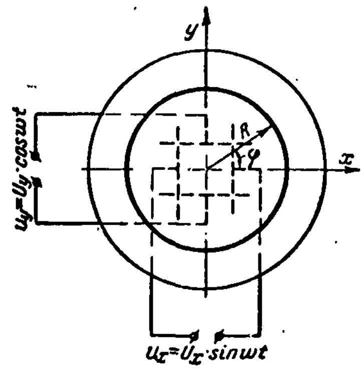

Fig. 15-20. Circular electron-beam sweep (the dashed lines show the arrangement of the deflecting plates).

b) Straight-line sweep over a raster. This widely employed type of scanning is obtained by applying sawtooth voltages with different periods to both pairs of deflection plates (or coils). As a result, the screen will show a representation of the voltage having the shorter period, in the form of a series of parallel bright lines. These lines will form a rectangle, in the absence of deflection distortions, a rectangle whose sides are defined by the amplitudes of the sweep voltages (Fig. 15-19).

The screen will show a trace of the beam as it moves with a gradual change in the sweep voltage. The flyback of the beam will cause almost no brightening of the screen, owing to its high speed as it moves over the screen. In the majority of cases, voltage pulses, synchronized with the sweep voltage, are applied to the tube modulator, so that the tube will be cut off during the beam flyback time.

The number of lines is equal to the ratio of the sweep-voltage periods. The rectangle formed by the lines is called a rectangular raster; its height determined by the amplitude of the sweep voltage

with the greater period (vertical-sweep voltage), is called the vertical height, while the width, which depends upon the amplitude of the voltage with the shorter period (the horizontal-sweep voltage), is called the line width. The position of the bright spot in the raster at any moment of time is determined by the instantaneous values of both sweep voltages.

c) <u>Circular electron-beam scanning and radial deflection systems</u>. If sinusoidal voltages, 90° out of phase, are applied to both pairs of deflection plates, where the maximum deflections of the beam are equal for both pairs of plates, the beam will describe a circle on the screen. Actually, if a voltage $u_x = U_x \sin \omega t$ is applied to the plates sweeping the beam along the X axis, and a voltage $u_y = U_y \sin(\omega t + 90^{\circ}) = U_y \cos \omega t$ is applied to the second pair of plates (Fig. 15-20), the corresponding deflections at any instant of time t equals:

$$x = k_x U_x \sin \omega t \tag{15-16}$$

and

$$y = k_y U_y \cos \omega t. \tag{15-17}$$

Squaring the right and left sides of Eqs. (15-16) and (15-17) and adding:

$$x^2 + y^2 = k_x^2 U_x^2 \sin^2 \omega t + k_y^2 U_y^2 \cos^2 \omega t. \tag{15-18}$$

If $k_x U_x = k_y U_y = R$, which can always be done by selection of the amplitudes of the sweep voltages, then Eq. (15-18), connecting the x and y coordinates of the position of the bright spot on the screen, may be written as follows:

$$x^2 + y^2 = R^2 (\sin^2 \omega t + \cos^2 \omega t) = R^2. \tag{15-19}$$

Equation (15-19) is the equation of a circle with a radius R, whose center lies at the center of the screen. The position of the bright spot on the screen at each moment of time is determined by the radius of the swept circle and the angle φ. Clearly, the angle φ is proportional to the time t and, consequently, represents the time in the scale of frequency of the sweep ω. If in addition to the sweep voltage, the voltage from any other electric signal is applied to one of the pairs of deflecting plates, the beam will be deflected from the sweep circle; the change in the angle φ will be an indication of the way in which the applied signal depends on time.

In the case described, the deflection of the beam from the circle occurs in the direction of one of the coordinate axes. It is more convenient, however, to study electric signals with a radial deviation of the electron beam. In order to obtain radial deflection with a circular sweep, special deflection systems are used.

One possible tube design for a radial-deflection system is shown in Fig. 15-21. Here, in addition to the deflecting plates, the deflection system contains radial-deflection electrodes located following the deflection plates, and taking the form of two coaxial hollow truncated cones. The electric field formed between the surfaces of the inner and outer cones deflects the beam in a radial direction at all times. Circular scanning is set up in such a tube by the method described above, and the signal setting up radial beam deflection is applied to the radial-deflection electrodes. With such a system, the magnitude of the radial deflection is connected nonlinearly with the magnitude of the signal voltage.

Radial beam deflection may also be obtained with the aid of another simpler deflection-electrode design (Fig. 15-22). The radial-deflection electrode takes the form of a pin sealed into the center of the tube

screen. The signal voltage that deflects the beam in the radial direction is applied between the pin and a conducting coating that is connected with the anode of the tube.

The beam, which is swept in a circle with the aid of the deflect-

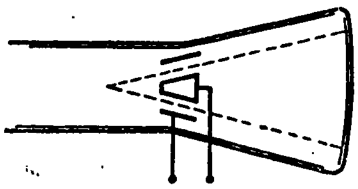

Fig. 15-21. Diagramatic representation of the arrangement of a tube for radial deflection, accomplished with the aid of two coaxial conical electrodes.

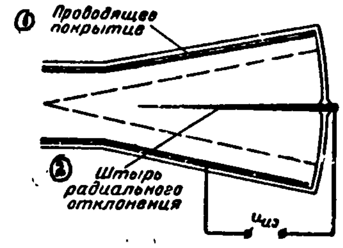

Fig. 15-22. Tube with central pin.
1) Conducting coating;
2) radial-deflection pin.

ing plates, is deflected by the field between the pin and the conducting coating. A drawback to this design in comparison with the conical electrodes for radial deflection is the low sensitivity of the pin system, caused by the great distance between the central pin and the coating.

A special case of a circular sweep with radial deflection is the case in which the beam is swept over a radial-circular raster. Such a sweep may be obtained, for example, with the aid of a single pair of deflecting coils rotating about the axis of the tube. If the current in such coils is constant, the light spot will move along a circle with an angular velocity equal to the angular velocity of coil rotation. If the current in the rotating coils changes periodically (for example, if the current in the deflecting coils is a sawtooth), the maximum and minimum values of deflecting current will correspond to the bright spot being positioned respectively in the center of the screen and at the greatest distance from it, so that the bright spot will be simultaneously

shifted along a radius and about the center. If the period of rotation of the spot is several times less than the period of the change in the deflecting current, the electron beam will be swept along a spiral. Where the deflecting current changes with a small period in comparison with the period of rotation, the raster created will appear as a bright radius rotating about the center of the screen.

The types of electron-beam scanning described above are very widely utilized in practice, but do not exhaust all possible types of sweeps.

15-5. CATHODE-RAY TUBE SCREENS.

Many materials begin to give off light under the influence of electron bombardment. This phenomenon is known as cathode luminescence, and materials that glow intensely under electron bombardment are called cathode phosphors.* A layer of phosphor applied to the glass end of a cathode-ray tube envelope forms the screen.

The light from cathode-ray tube screens is observed on the side opposite to that which is bombarded by the electrons. Clearly, the thickness of the layer of phosphor should be such that, first, it is greater than the depth of penetration of the electrons, since if this were not the case, the electrons of the beam would strike the glass of the envelope, giving up their energy to it and, second, the layer should be sufficiently thin that the glow excited by the electrons is not strongly absorbed as the light passes through the layer. The depth of penetration of electrons into a layer of phosphor depends upon the energy of the electrons and the properties of the phosphor. Thus, the

---

* Phosphors are defined as substances that may be excited so as to give off light by various external energy sources (ultraviolet radiation, X rays, etc.). In this case, we are concerned solely with cathode-ray phosphors.

thickness of the phosphor layer is selected in each individual case as a function of the type of tube and its operating regime.

The basic specifications for a phosphor are:

1. The phosphor should efficiently convert the energy of the electrons into luminous energy.

2. The phosphor should give light of the required color.

3. The phosphor should have good vacuum properties, i.e., should volatilize little in a vacuum, should be sufficiently heat resistant, and capable of good outgassing.

4. The useful life and stability of phosphor properties should provide for normal operation of the cathode-ray tube for the established service life.

In addition to the requirements enumerated, the most important characteristic of phosphors is their persistence, i.e., the time during which the phosphor still glows after electrons have ceased to act upon it. As a rule, the persistence is assumed to be the time required for the brightness to drop (after excitation has ceased) to one per cent of the initial value.

For the majority of phosphors, their basic properties are determined by the presence of admixtures that are called activators. These activators influence the conversion efficiency with which electron energy is converted into light energy, the color given off by the phosphor, and the persistence. By changing the amount of activator in a phosphor, or the type of activator, it is possible to change the properties of the phosphor.

The basic materials used for phosphors are as follows:

1) Sulfides, i.e., compounds of certain materials with sulphur, and mixtures of these compounds;

2) silicates, i.e., salts of silicic acid and their mixtures;

3) tungstates, i.e., salts of tungstic acid.

In addition, zinc oxide, selenides, and many other materials may be used as phosphors. Admixtures of metals are normally used as activators. The most common activators are silver, manganese, and copper.

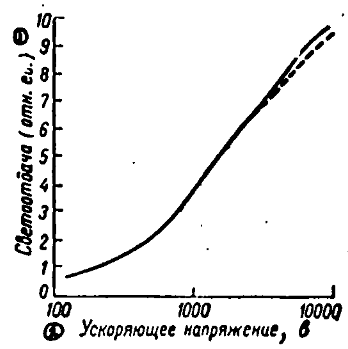

Fig. 15-23. Sample curve showing the dependence of light yield on accelerating voltage.
1) Relative units; 2) accelerating voltage, v.

One of the most important parameters of a phosphor screen is its light output, determining the efficiency with which electron energy is converted into light energy. Light output is determined by the candle power delivered by the screen per watt of power attributable to bombardment by electrons, and is normally measured by candles per watt (candles/watt). The light output, which depends upon the properties of the phosphor used for the screen, also in great degree depends upon the structure of the screen and the method used to apply the phosphor. The fact of the matter is that a portion of the luminous flux produced by the phosphor when excited by electron bombardment may be absorbed in the phosphor layer and in the glass of the screen, reflected from the surface of the glass, etc. Thus, in order to specify the magnitude of light output, it is necessary to specify the conditions under which the screen is to be used.

The candle power delivered by a screen equals the product of the light output by the power of the electron beam:

$$I_{св} = LP_{л}, \; * \tag{15-20}$$

where L is the light output; $P_{л}$ is the power of the beam, equal to the product of the beam current $i_{л}$ by the magnitude of the potential difference through which the electrons have fallen (the accelerating voltage).

Consequently, when the beam current or the accelerating voltage is increased, there should be a proportional rise in the screen candle power. Since the magnitude of the light output itself does not remain constant, however, when the accelerating voltage and beam current change, the relationship between candle power and beam power turns out to be nonlinear.

The light output of screen phosphors is strongly dependent upon electron-beam energy: as the accelerating voltage rises, light output increases; the dependence of light output upon accelerating voltage, in this case, is nonlinear for the majority of phosphors. Figure 15-23 shows a sample curve representing light output as a function of accelerating voltage.

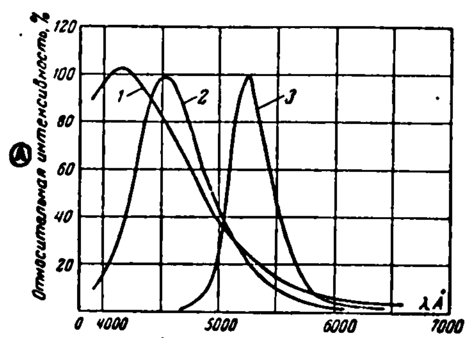

Fig. 15-24. Sample spectral characteristics of screens.
1) Type M screen ($CaWO_4$); type A screen (ZnS:Ag); 3) type I screen (Zn_2SiO:Mn); A) relative intensity, per cent.

---

* [$I_{с в}$ - I_{sv} - I_{cvet} - I_{light}; $P_{л}$ - P_{luch} - P_{beam}.]

For small beam-current densities (up to 1-2 $\mu a/cm^2$), the magnitude of the light output is independent of the current density, and the candle power rises in proportion to the beam current. For large current densities, the light output drops somewhat as current density rises and, consequently, screen candle power rises more slowly than does beam current. This phenomenon is called current saturation.

In practice, the electron beam is continuously shifted over some portion of the surface of the tube screen with a predetermined velocity. The visibility of the electron-beam track is in this case determined not by the candle power delivered by the entire surface of the screen that has been subjected to electron excitation, but by the brightness, which is determined by the ratio of the total candle power to the area of the glowing surface. The greater the surface area excited by the electron beam per unit time, the greater must be the beam power in order to obtain the required brightness (given constant light output).

A second important characteristic of screens is the spectral composition of the light. The distribution of radiant energy of phosphors (intensity of luminescence) over the wave-length spectrum is represented by the spectral characteristics of the phosphors. Figure 15-24 gives examples of such spectral characteristics for certain screens. It is easily seen that phosphors have a well defined maximum intensity of luminescence. The wavelength of the light that corresponds to maximum luminescence intensity is one of the parameters of a phosphor. In some cases, it is possible to shift the spectral characteristics in the required direction by proper choice of filters located in front of the screen.

Where the color required differs from the color of phosphors in use, a mixture of phosphors is applied to the screen. Thus, for example, for the screens of television picture tubes which have a white lumines-

cence, a mixture of phosphors is used whose components taken separately give blue and yellow luminescences. The resultant spectral characteristic of such a screen is shown in Fig. 15-25.

The double layer or "cascade" screens form a separate group. A

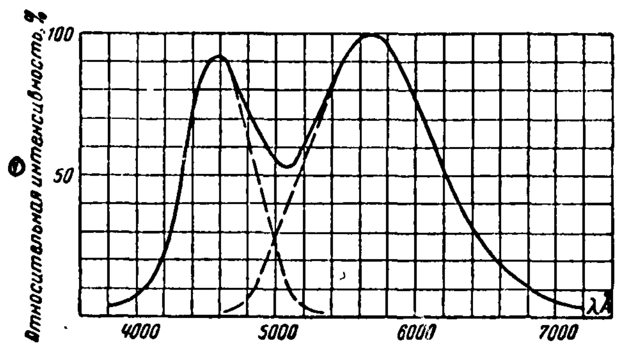

Fig. 15-25. Spectral characteristics of two-component B screen. The dashed line shows the spectral characteristics of the component phosphors.
1) Relative intensity, per cent.

screen of this type takes the form of two layers of phosphor applied to glass. The upper layer of phosphor, facing the electron gun, is excited by electron bombardment. The ultraviolet portion of the luminescence of this phosphor in turn excites the second layer of phosphor (applied to the glass). The basic property of this type of screen is the long persistence of the phosphors excited by ultraviolet radiation rather than by electron bombardment. An example of a cascade-type screen is the screen that consists of a layer of a mixture of zinc and cadmium sulphide, activated by copper, which gives a yellow light, and a layer of zinc sulphide, activated by silver, applied over the first layer, and giving a blue light. Under the action of the electron beam, the second layer radiates blue light and simultaneously excites the first layer which gives yellow light. As a result, the resultant screen luminescence is white. When electron excitation ceases, the blue component is eliminated, and a yellow after-glow remains.

TABLE 15-1.
Properties of certain cathode-ray tube screens.

1 Обозначения	2 Характеристика слоя	3 Состав люминофора	4 Цвет свечения	5 Максимум спектральной характеристики, Å	6 Время послесвечения	7 Основное применение
А	8 Однослойный	$ZnS:Ag$	13 Синий	4 500	18 Короткое	23 Осциллографические трубки
Б	9 Однослойный, двухкомпонентный	$ZnS:Ag$ и $ZnS \cdot CdS:Ag$	14 Белый	4 600 и 5 700	19 Короткое	24 Кинескопы для непосредственного наблюдения
В	10 Двухслойный, каскадный	$ZnS:Ag$ на $ZnS \cdot CdS:Cu$	15 Белый	4 400 и 5 600	20 Длительное	25 Индикаторные трубки
И	11 Однослойный	$Zn_2SiO_4:Mn$	16 Зеленый	5 200	21 Среднее	26 Осциллографические и индикаторные трубки
М	12 Однослойный	$CaWO_4$	17 Голубой	4 200	22 Короткое	27 Осциллографические трубки

1) Designation; 2) characteristic of layer; 3) composition of phosphor; 4) color; 5) maximum in spectral curve, A; 6) persistence; 7) basic application; 8) single layer; 9) single layer-double component; 10) two-layer, cascade; 11) single layer; 12) single layer; 13) blue; 14) white; 15) white; 16) green; 17) azure blue; 18) short; 19) short; 20) long; 21) medium; 22) short; 23) oscilloscope tube; 24) direct-viewing picture tubes; 25) indicator tube; 26) oscilloscope and indicator tubes; 27) oscilloscope tubes.

On the basis of persistence, all screens are arbitrarily divided into five groups: very short-persistence phosphors — persistence less than 10^{-5} sec, short-persistence — from 10^{-5} to 0.01 sec, medium-persistence — from 0.01 to 0.1 sec, long-persistence — from 0.1 to 16 sec, and very-long-persistence, above 16 sec. The choice of a phosphor with respect to persistence is determined by the function of the tube.

Basic data for the most common phosphors are given in Table 15-1.

The electrical resistance of phosphors is very high, and for all practical purposes they may be assumed to be insulators. Thus, electrons arriving at the screen are carried off solely as a result of secondary emission from the screen, occurring as a result of the electron bombardment. The secondary electrons emitted from the surface of

the screen move toward the anode of the tube; in order to improve conditions for removing secondary electrons, a conducting coating (normally aquadag) is applied to the inside surface of the tube envelope.

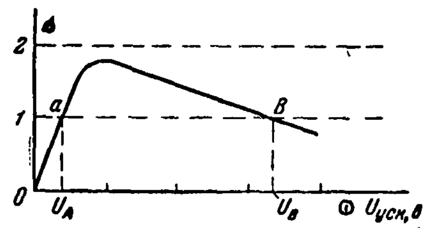

Fig. 15-26. Screen secondary-emission coefficient as a function of the voltage accelerating the primary electrons.
1) U_{usk}, v.

The secondary-emission coefficient of insulators, as well as for conducting surfaces, depends upon the energy (velocity) of the primary electrons. Figure 15-26 gives a sample curve illustrating the dependence of the secondary-emission coefficient of an insulator upon the accelerating voltage, which determines the energy of the primary electrons. As the primary-electron energy increases, the secondary-emission coefficient rapidly rises, reaching a maximum at a primary-electron energy on the order of 300-800 ev, and then begins to drop slowly. The secondary-emission coefficient becomes equal to one at two values of accelerating voltage: at point A and at point B (Fig. 15-26).

If the secondary-emission coefficient of a screen does not equal one for an accelerating-voltage value, the screen will store charges until its potential becomes such that the velocity of the electrons arriving at the screen either corresponds to $\sigma = 1$, or turns out to equal zero, i.e., the total screen current becomes zero. When this happens, further storage of charges on the screen ceases. Thus, where the accelerating voltage on the anode of the tube lies between zero and U_A, the secondary-emission coefficient is less than one and, consequent-

ly, the number of incoming electrons is greater than the number of secondary electrons. Negative charges begin to be stored on the screen until the difference in potential between the tube anode and the screen becomes equal to the potential difference between the anode and the cathode, i.e., until the screen potential becomes equal to the cathode potential. When this happens, the electrons lose all of their velocity on the path from anode to screen, and screen current ceases. For anode voltages from U_A to U_B, the secondary-emission coefficient becomes greater than one, and the screen becomes positively charged until its potential becomes comparable with the anode potential, or turns out to be 1-2 volts greater. When this happens, the excess secondary electrons return to the screen, and an equality is set up between the number of electrons arriving at the screen and the number traveling from the screen to the anode. In this case, the energy of the electrons arriving at the screen and, consequently, the brightness of the screen as well, will be determined by the anode voltage.

If the anode voltage of the tube becomes greater than U_B, and the secondary-emission coefficient is less than one, as a result of the storage of charges on the screen, its potential becomes equal to U_B, and the electrons in the beam are retarded as they travel from anode to screen to a velocity corresponding to this potential. Consequently, a further increase in anode voltage will not increase the energy of the electrons incident upon the screen or the brightness of the screen. The magnitude of the potential U_B, which depends upon the properties of the phosphor and which is called the maximum phosphor potential, limits the possible increase in screen brightness that can result from an increase in accelerating voltage.

The maximum potential for tube screens differs from the maximum potential of the phosphor employed, since in the formation of secondary

emission, the surface of the glass of the screen is involved, and in addition, the shape of the secondary-emission characteristic may change as a result of the incidence upon the phosphor surface of particles of getter and various contaminants. The maximum potential changes over the useful service life of the tube, in addition, owing to changes in the properties of the phosphor under electron and ion bombardment, gas evolution from parts of the tube, etc. As a rule, the maximum screen potential drops over the useful service life, and brightness gradually decreases.

In order to eliminate the effect of the maximum screen potential, which is especially important for tubes with high accelerating voltages (greater than 8-10 kv), a conducting coating is applied to the inside surface of the screen in the form of a thin metal film which is quite transparent to the electron beam. A material very commonly used to form such a film is aluminum, applied to the surface of a screen by evaporating it by rapid heating in the evacuated envelope, on the bottom of which the phosphor has first been applied. The thickness of such an aluminum film varies from 0.5 to 3 microns and more, depending upon the accelerating voltage. The metal film is connected with the tube anode and, consequently, the screen is always at a potential equal to the anode potential.

The advantage of a screen with a thin metal film lies not only in the fact that the effect of the maximum phosphor potential is neutralized, but also lies in an increase in the light output owing to reflection from this layer of radiation directed toward the interior of the tube. There is some loss in light output owing to a partial loss of energy of electrons in the metal film, but this is compensated by an increase in the light from the radiation reflected from the film. The utilization of such a reflecting metal film may yield an increase in

light output (for thin films) of up to 20-25 per cent.

In addition, the metal film, which protects the phosphor against damage, increases the service life of the screen. An important factor in screen damage is screen bombardment by negative ions produced by the oxide-coated cathode of the tube, for example, by ions of oxygen liberated as a result of electrolysis of the cathode oxide. Such ions, which possess high mass in comparison with electrons, strike the screen with great energy, and distort the crystals of the phosphor. In tubes with magnetic focusing and deflection, negative ions move toward the screen in a diverting beam, since the deflecting and focusing effect of a magnetic field is less the greater the mass of the moving charged particle, and the magnetic field, focusing and deflecting electrons, has practically no effect upon the path of motion of the ion. As a result, a spot is formed in the center of the screen under the action of ion bombardment; this spot has a considerably lowered brightness, and is called an ion burn. In tubes with electrostatic focusing and deviation, the paths of the negative ions are similar to the paths of the electrons, and a gradual attack on the entire surface of the screen results from ion bombardment.

The metal film protects the phosphor against ion bombardment. The film thickness needed for this depends upon the accelerating voltages. In some cases, protection against the appearance of an ion burn is possible only where the film is so thick that the light output of the screen drops by 15-20 per cent owing to the energy lost by the electrons in the film.

15-6. OSCILLOSCOPE AND RADAR-INDICATOR CATHODE-RAY TUBES.

As we have indicated above, the position of the bright spot on the screen of a cathode-ray tube is determined at any moment of time by the

instantaneous values of the deflecting voltages. On the other hand, the rate of motion of the bright spot over the screen depends upon the rate of change of the deflecting voltages. This property of cathode-ray tubes makes it possible to utilize them as measuring devices in investigations of changing electrical processes. A tube with a linear electron-beam sweep may be used as such a measuring device. If to one pair of plates the periodic voltage to be measured is applied, and to the second a sawtooth voltage is applied whose amplitude and period are known, on the basis of the scanning, it is possible to obtain comprehensive information on the measured voltage. The direction in which the beam is swept by the sawtooth voltage may be taken as a time axis, since the deflection of the bright spot in this direction is a linear function of time (within a single cycle), while the deflection in the perpendicular direction indicates the instantaneous value of the measured voltage as scaled by the sensitivity. Thus, it is possible to determine the frequency, amplitude, and nature of change in a measured periodic voltage.

In like manner, it is possible to study nonperiodic or transient (one-time) processes. Since in this case the motion of the bright spot over the screen does not repeat itself and, consequently, a stable representation of the measured voltage does not appear on the screen, in order to investigate such processes it is necessary to use a tube having long screen-persistence, or to photograph the trace of the bright spot.

Cathode-ray tubes designed to record alternating processes converted to electric signals, visually or by means of a photographic record, are called respectively oscilloscope or oscillograph tubes.* The region of application of the tubes also determines the requirements

---

* From now on, we shall refer to oscilloscope and oscillograph tubes as oscillograph tubes.

applying to tube parameters.

An important characteristic of an oscillograph tube is its resolving power, determined by the width of the bright line. The width of the bright line basically depends upon the quality of the tube focusing system. It is very important that the width of the bright line near the center of the tube differ little from its width near the edges. A change in line width upon deflection is caused by beam defocusing, the causes of which were considered above. The line width at a specific distance from the center of the screen is frequently taken as a parameter for electrostatic-deflection tubes. For tubes with magnetic deflection, only the line width at the center is taken into consideration, since the degree of beam defocusing upon deflection depends strongly upon the external deflecting system. As has already been said, with magnetic deflection, beam defocusing is easily decreased and, consequently, bright-line width is easily made smaller.

Certain parameters of oscillograph tubes are connected with screen luminescence. Among these parameters are screen brightness at a given value of beam current, luminescence color for the spectral characteristic, the color and persistence of afterglow.

Where the beam current remains constant, brightness depends upon screen properties and the energy of the electrons in the beam, which is determined by the accelerating voltage. In many cases, in order to obtain the required brightness, it proves necessary to increase the accelerating voltage. Thus, for example, this is necessary for tubes designed to record one-time high-speed processes. When the beam moves over the screen at very high rates of speed, brightness decreases, since the mean density of the current on the screen proves to be very small. An increase in electron energy, however, owing to an increase in voltage on the electron-gun anode leads to a decrease in sensitivity.

In cases where it is necessary to increase electron-beam energy with no loss of sensitivity in deflection, an additional acceleration of beam electrons is used after they have passed through the deflecting

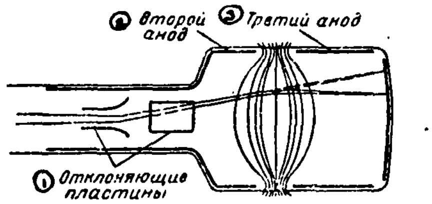

Fig. 15-27. Refraction of deflected beam by the field of the lens formed between the second and third anodes. 1) Deflection plates; 2) second anode; 3) third anode.

plates; this is called after-acceleration. For this purpose, rings of a conducting coating (normally aquadag) are applied to the inside walls of the wider section of the tube envelope; the rings are insulated from the conducting coating that is connected to the second anode of the tube. A voltage considerably higher than the second-anode voltage (by a factor of 1.5-2, approximately) is applied to the conducting rings. In order to avoid arcing between the after-acceleration rings and the conducting coating that is connected to the second anode, the surface of the glass between them is covered with a semiconductor having high resistivity (normally chromium oxide or iron oxide).

Under the action of the accelerating field of this coating, which is a third anode of the tube, the electrons, which have already been deflected, receive additional energy on their path to the screen.

The after-acceleration electrode (third anode), together with the conducting coating that is connected with the second anode of the electron gun, forms an electrostatic lens (Fig. 15-27). This lens, refracting the electron beam, shifts it somewhat toward the center of the screen, owing to which the deflection sensitivity is somewhat de-

creased. In addition, the lens produces additional distortion in beam deflection. Clearly, the effect of this lens on the position of the bright spot on the screen will be less for small deflections of the electron beam by the deflecting plates than for large deflections.

In order to eliminate the distortion introduced by after acceleration, multiple after acceleration is utilized, set up by a series of conducting rings whose potential is gradually increased to a maximum potential equal to the required third-anode potential. With relatively small changes in potential from ring to ring, the refracting action of the lens formed by them turns out to be small. The widened section of the tube envelope, with the after-acceleration, is made cylindrical in shape, which increases the diameter of this lens, and, consequently, decreases field nonuniformity in the region where the beam passes through the lens.

The choice of luminescence parameters is determined by tube function. Thus, in tubes intended for visual observation of alternating electric processes without photography, screens with green or orange luminescence are normally used, while in tubes which are to be photographed, screens with blue or azure-blue luminescence are used, since they have the maximum effect upon the light-sensitive emulsion.

The choice of persistence is dictated by the nature of the electric processes being studied. In studying periodic, slowly varying processes, screens are utilized that have short and medium persistence, while tubes for visual investigation of nonperiodic or transient processes use long-persistence screens. In this case, it is as if the beam trace whose motion corresponds to the rapidly occurring process were written on the screen of the tube.

A special type of tube with controlled persistence is the dark-trace tube – the skiatron. In these tubes, instead of phosphors, a

material is applied to the screens which changes color under electron bombardment; such a material is potassium chloride (KCl), whose crystals turn violet for a long period of time under the action of an electron beam (when the screen is illuminated with white light).

To remove an image from such a screen, it is necessary to heat it or to subject it to an electric field; the first method is the more common. The rate of decrease in image contrast depends upon the heating temperature. Heating is carried out either by means of an external source of heat, or by passing an electric current through a transparent metal film which serves as a backing for the layer of potassium chloride. In this case, the screen is applied not to the glass at the end of the envelope, but to a mica plate that is first coated with the metal film. By decreasing the heating current through the backing, it is possible to vary screen persistence from comparatively small values (several seconds) to hours and days. Such tubes are designed to record non-periodic, not-too-rapid processes. Thus, for example, the photograph of the reverse side of the moon taken by the first Soviet automatic interplanetary station in 1959 was recorded on skiatron screens.

Parameters of cathode-modulator unit of a tube. Among such parameters are the cutoff voltage on the modulator and the amount of modulation.

The cutoff voltage, which under given conditions depends upon the relative arrangement of cathode, modulator, and accelerating electrode, determines the required power-supply voltage.

In this case, by modulation we mean the difference in value between the cutoff voltage and the voltage on the modulator that corresponds to the nominal value of beam current. Modulation characterizes the controlling action of the modulator on the magnitude of the beam current, and depends upon the geometric dimensions of the electron-gun

and upon the emission properties of the cathode. For oscillograph tubes, the modulation is an auxiliary parameter characterizing cathode quality.

Deflection-system parameters. The deflection sensitivity is the basic deflection-system parameter. This parameter is meaningful for tubes with electrostatic deflection, since deflection sensitivity in tubes with magnetic deflection to a large degree depends upon the structure and parameters of the deflection system. The attainable sensitivity of a tube with magnetic deflection is determined by the construction of the tube itself, i.e., the accelerating voltage and tube geometry (distance from neck of tube to screen, diameter of neck and distance from electron gun to beginning of conical section of envelope).

Tube auxiliary parameters. The auxiliary parameters include quantities that determine possible operating regimes of tubes and circuits. Such parameters are the insulation resistance between tube electrodes (or the magnitude of leakage currents between electrodes in a given regime), and interelectrode capacitances.

In addition, for tubes with electrostatic focusing, the limits of variation in first-anode voltage within which the best beam focusing is obtained should be specified.

Tube service life. Tube failure, as a rule, is connected either with destruction of the screen, or with a deterioration in cathode emission properties (especially in the center of the oxide-coated section). A criterion of lifetime is normally the change in screen brightness under given tube operating conditions. In addition, the width of the bright line may be used as a criterion of lifetime. A change in width upon deterioration in cathode emission is caused by the fact that, in this case, it is necessary to increase the modulator voltage to obtain the required beam current. This results in an increase in the size of the (emitting)

surface of the cathode used and, correspondingly, in the size of the bright spot.

In addition to the oscillograph cathode-ray tubes described, there

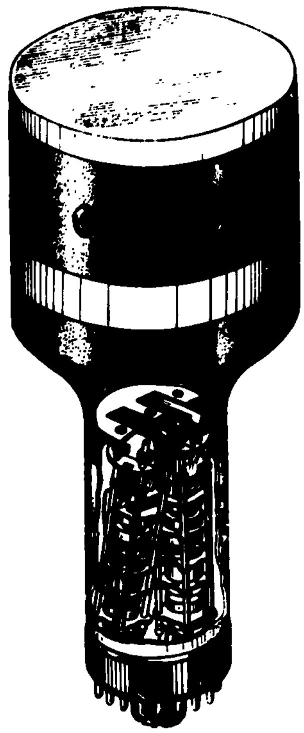

Fig. 15-28. Example of construction of two-beam oscillograph tube.

are others as well, utilized in various special measuring devices. Of these designs, let us take note of the double-trace oscilloscope tubes.

A double-trace oscilloscope tube (Fig. 15-28) is designed for comparative study, at the same time, of two alternating electric processes. The tube has two electrooptical systems, whose optical axes meet at the center of the screen. In the design shown in Fig. 15-28, the second anodes of both optical systems are connected, while the first anodes, modulators, and deflection plates are supplied separately.

Generally speaking, it is possible to construct oscillograph tubes with more than two beams. Such multibeam designs are utilized in special measuring instruments that record several alternating processes

61-102

simultaneously.

Cathode-ray tubes are widely utilized as radar indicators.

The probing pulse sent out by the radar transmitter and the echo pulse returning after reflection are applied to a single pair of deflecting plates in a cathode-ray tube, while a sawtooth sweep is applied to the second pair; the sweep voltage is equal in period to the probing-pulse signals. The position of the probing pulse on the screen of the indicator tube remains unchanged, while the location of the echo pulse along the sweep line depends upon the distance to the object observed. As we have said above, sweeping the electron beam with a sawtooth voltage forms a time axis; the length of the line drawn on the screen by the bright spot as a result of the action of the sweep voltage reflects the period of the sweep voltage in the scale. This axis may be graduated in distance units, and the distance to the observed object may be measured from the screen. Such a device is called a range indicator.

Either standard oscilloscope tubes may be used in range indicators, or tubes with circular sweeps and radial deflection, which permit a more accurate measurement of distance since in such tubes the time-sweep line is longer.

In addition to indicators in which the echo pulse deviates the bright spot on the screen (indicators with amplitude marking), indicators are used that have brightness marking, in which the echo pulse is applied to the tube modulator, which produces a point whose brightness is considerably greater than the brightness of the screen as a whole at the instant that the echo pulse arrives at the screen. In such indicators, two sawtooth sweep voltages are applied to the deflection plates or coils; this forms on the screen a raster consisting of a group of horizontal or vertical lines. The receiving device for such a radar

installation has a directional antenna that rotates about its vertical axis. The period of one of the sweep voltages equals the time required for one revolution of the antenna. Consequently, the shift in the very bright point in the direction of the sweep by this voltage corresponds to the angle of rotation of the antenna at which the reflected signal was applied to the receiver. The second sweep voltage may have a period equal to the period of the signal pulses; in this case, the initial instant of the sweep coincides with the moment at which the probe pulse is sent. A shift in the bright point in the direction of the second sweep voltage determines the time interval between the sending of the probe pulse and the arrival of the echo pulse, i.e., the range of the object.

The angle of rotation of the antenna about its vertical axis, measured with respect to the North direction, is called the azimuth of the object. Thus, on the basis of the position of the bright spot on the screen of the indicator, it is possible to determine the two space coordinates of an object: range and azimuth. The third coordinate of the object is the elevation, i.e., the angle between the direction of the object and the horizontal. By synchronizing one of the sweep voltages with the rotation of the radio beam transmitted into space about the horizontal axis, it is possible to determine the elevation on the basis of the position of the bright spot on the screen. As a result, cathode-ray display tubes may be used to measure all coordinates specifying the position of an object in space; here, a single indicator can specify no more than two coordinates. Plan-position-indicator type radar installations are in wide use; in these units, the indicators use sweeps generating a radial-circular raster. On the screen of such a tube, we can observe what amounts to an outline map of the locality, on which bright blips indicate the position of reflecting objects. The structure of the tubes depends upon their functions. The majority of

oscillograph tubes designed for measuring devices use electrostatic focusing and deflecting systems. The advantages of an electrostatic system are small power drain (in comparison with a magnetic system), and a less cumbersome installation for the tube. In addition, electrostatic deflection provides a linear relationship between the shift in the beam and the deflection voltage, which is important for instruments used to measure the amplitude of radiated voltages. Cathode-ray indicator tubes may use either electrostatic or magnetic focusing and deflection systems. In order to obtain a circular sweep with radial deflection, systems are frequently utilized in which the deflecting coils rotate about an axis coinciding with the axis of the tube, which causes the electron beam to rotate.

15-7. TELEVISION PICTURE TUBES.

In 1907, the Russian scientist Professor B. L. Rozing conceived the idea of utilizing cathode-ray tubes for reproducing pictures. Practical utilization of his suggestion, however, proved to be possible only after about twenty years.

The principle of transmitting and reproducing television images is based upon breaking down the transmitted image into a large number of elements. To each element of the image there corresponds a value of brightness. In the transmitting device, the alternation of picture-element brightnesses is converted into a series of successive electric signals (videosignals) that modulate the carrier frequency of the transmitting station. In the television receiving device, the received videosignals are applied to the modulator of a cathode-ray tube, and as a result, the beam current and, consequently, screen brightness at the point of incidence of the beam, varies in accordance with the change in videosignal voltage.

Two basic conditions must be fulfilled for the reproduction of

pictures on the screen of a receiving tube; first, the position of the bright spot on the screen of the tube at any instant of time should correspond to the position of that element in the transmitted image from which the signal applied at that instant to the modulator has been taken and, second, all elements of the image should be reproduced in a time less than the time dictated by the persistence of light perception by the human eye. Since perceived light is retained by the eye for a period of about 0.1 sec, the entire sequence of picture elements should be reproduced in a somewhat smaller time interval, which will provide for simultaneous reproduction of the entire picture.

Clearly, the greater the number of elements into which the image is broken down, the more clearly it may be reproduced. Normally, the alternation of elements is chosen in terms of raster horizontal lines. Motion of the beam over the screen should be synchronized with the conversion of the bright elements of the image into electric signals in the transmitting device. In modern transmitting television installations, the conversion of light into electric signals is also accomplished with the aid of an electron beam moving over a projected image. Consequently, the motion of an electron beam in a picture tube is synchronized with the motion of the electron beam in the transmitting tube; the operating principles of transmitting tubes are discussed below.

In order to sweep the electron beam along the lines, sawtooth voltages are applied to both pairs of deflection coils (or plates); in this case, the number of lines is determined by the ratio of the periods of these voltages. The period of the vertical sweep is so chosen that the time required to sweep from top to bottom of the raster will be less than 0.1 sec. As a rule, the frequency of the vertical-sweep voltage is set equal to the AC line-voltage frequency (50 cps). In this case, the beam will cover the entire frame in 0.02 sec.

According to the standard adopted in the USSR, the number of lines is set at 625. Such a division of the image provides high-quality reproduction. Clearly, in order to avoid having a dark band visible on the screen between the lines, the width of the bright line should be approximately equal to the vertical frame height divided by the number of lines. On the other hand, if the width of the line is increased beyond this value, it will lead to overlapping of the lines, and a deterioration in picture sharpness.

Tubes utilized for the reproduction of pictures are called kinescopes. For a kinescope to provide high-quality reproduction of pictures, it should possess several properties which are determined basically by the following parameters and characteristics:

1. The resolving power of the kinescope depends upon line width and determines the sharpness of the image. In turn, the line width required for a given number of lines depends upon frame size, i.e., the size of the picture reproduced on the screen. The ratio of the size of the frame is normally taken equal to 4/3. Where the height of the frame is increased, there is a corresponding increase in line width.

For the picture to be sharp enough both at the center and at the edges, beam defocusing upon deflection should be small. Thus, in the majority of kinescopes, magnetic deflection is used, which provides less defocusing of the beam. In kinescopes with small picture dimensions, which require small line widths, electron guns with a second magnetic lens are utilized; this provides better beam focusing than does an electrostatic lens. For large picture sizes, it is possible to utilize electrostatic electron guns, since in this case, line widths may be relatively large. Thus, for example, in a kinescope with a picture measuring 135 x 100 mm, with 625 lines, the line width should equal about 0.15 mm, while in a kinescope with a picture measuring 370 x 280 mm,

line width is on the order of 0.5 mm. The utilization of electron guns with electrostatic focusing in kinescopes with large pictures makes television receivers cheaper, since there is no need for cumbersome focusing and power-supply circuits.

2. Screen luminescence, determined by brightness, color, and persistence. The required brightness depends upon the conditions under which the image is observed. For good tone gradations to be visible on the reproduced image under normal room illumination conditions, a screen brightness of from 3 to 100 millistilbs is sufficient.

The basic factors determining screen brightness are beam current, accelerating voltage, and the properties of the screen phosphor. An increase in brightness owing to an increase in beam current (by increasing the constant voltage on the modulator) is possible only within certain limits, since in this case, the size of the cathode emitting surface is increased and, consequently, the size of the bright spot will also increase, which decreases picture sharpness. Thus, in the majority of kinescopes, the current from the cathode is about the same and does not normally exceed 100-150 μa. The required screen brightness is maintained when picture size is increased by increasing the accelerating voltage. If in kinescopes with pictures measuring 135 x 100 mm, an accelerating voltage on the order of 3-5 kv is required (for a brightness of about 3 millistilbs), then for kinescopes with pictures measuring 320 x 240 mm, the accelerating voltage will range from 12-14 kv.

The best color for television screens is white, which provides a black-and-white picture. Thus, a two-component phosphor is generally used for the screens of television tubes (type B screen). The persistence of the screen should be somewhat less than the time required to sweep a single frame, i.e., less than the period of the vertical-

sweep voltage.

3. Constrast of image on screen. Contrast is defined as the ratio of the brightness of the brightest picture element to the brightness of the darkest picture element. The picture may be assumed to be good

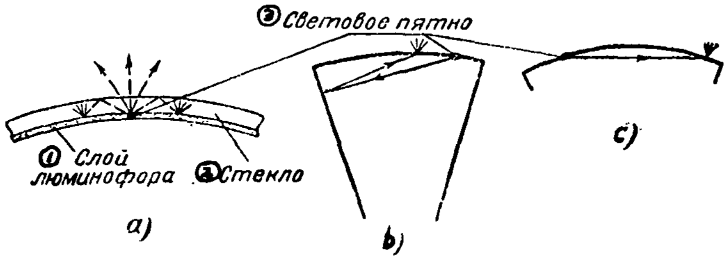

Fig. 15-29. Causes of loss of contrast.
a) Illumination of phosphor layer as a result of total internal reflection in glass at end; b) reflection of illumination from inner walls of envelope; c) direct illumination of unexcited portions of screen by radiation from the bright spot.
1) Phosphor layers; 2) glass; 3) bright spot.

enough where the contrast is on the order of 20-40. A decrease in contrast is caused by many factors that lead to a situation in which the radiation excited by the electron beam at any point on the screen will only pass through the end of the envelope to the viewer in part. A portion of this radiation will fall on other sections of the screen, owing to total internal reflection from the inside and outside surfaces of the glass, reflection from the inner walls of the envelope, and when there is curvature of the screen, some of this light will fall directly on other parts of the screen (Fig. 15-29). In addition, contrast is decreased owing to light from external sources that is reflected from the phosphor layer.

In order to reduce the effect of factors decreasing contrast it possible, first, to utilize smoked glass for the bottom of the envelope;

this type of glass is a better light absorber than standard glasses. Radiation incident upon the phosphor layer as a result of total internal reflection from the surface of the end of the tube, and light from external sources, are subject to multiple absorption in the glass, while radiation from the section of the screen excited by the electrons passes through the end of the envelope only once. The utilization of such a glass somewhat decreases screen brightness, but when this is done contrast increases several times.

Second, a considerable increase in contrast may be obtained by using aluminum-treated screens. An aluminum film that is transparent to the electron beam and opaque to light protects the screen against reflected light incident upon it from the walls of the envelope, and from direct illumination upon unexcited sections of the screen.

Combining both of these methods, it is possible to obtain a high value of contrast (on the order of 100). A further increase in contrast leads to no further improvement in the quality of the image that is noticeable to the eye.

4) Tube dimensions. The dimensions of television receivers are chiefly determined by the dimensions of the tube. The picture size chosen determines the minimum dimensions of the kinescope screen. With a circular screen, and a picture width-height ratio of 4/3, the area occupied by the picture amounts to about 60-65 per cent of the screen surface. Kinescopes with rectangular screens, in which the image occupies about 95 per cent of the screen surface, are considerably more economical with respect to television-set size. Correspondingly, the volume of a kinescope with a rectangular screen turns out to be considerably less than the volume of a picture tube having the same size picture, but a circular screen.

Decreasing the length of kinescopes, accomplished by shortening

the optical system of the tube and by decreasing the distance from the deflecting system to the screen, involves an increase in the total angle of deflection, which causes great complications, since in this case there is an increase in the defocusing of the electron beam. In modern

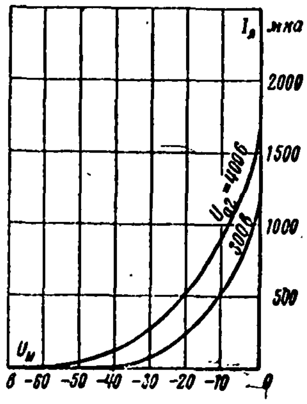

Fig. 15-30. Example of kinescope modulation characteristic.

kinescopes, the deflection angle (measured along the picture diagonal) reaches 110° in the best tube specimens.

5. <u>The modulation characteristic</u>, which determines the way in which the beam current depends upon the modulator voltage. For kinescopes, it is desirable to increase the slope of the modulation characteristic while simultaneously decreasing the absolute value of cutoff voltage; this makes it possible to utilize videosignals of lower amplitudes. Decreasing videosignal amplification decreases its distortion. Figure 15-30 shows a sample modulation characteristic for a kinescope.

Since magnetic deflection of the electron beam is utilized in the majority of kinescopes, the negative ions move to the screen, almost undeflected, in a somewhat diverting beam; as a result, an ion burn may appear in the center of the screen. In order to protect the screen against ion burn, kinescopes normally have so-called ion-trap systems, which lead the beam of ions off to the electron gun anode. The operating principle of ion traps is based upon the fact that a magnetic field

deflects a stream of electrons considerably more strongly than a stream of ions.

One of the most common electron-gun designs with an ion trap is shown in Fig. 15-31. The cathode-modulator unit, with a shielding elec-

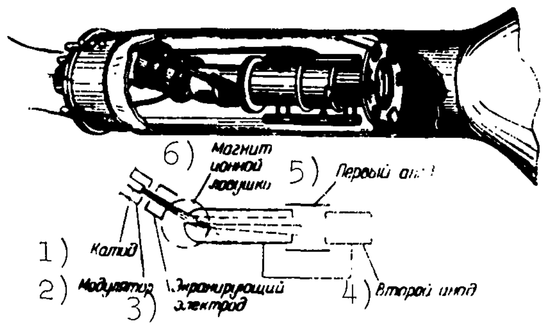

Fig. 15-31. Structure and operating principle of kinescope electron-gun with ion trap.
1) Cathode; 2) modulator; 3) screening electrode; 4) second anode; 5) first anode; 6) magnet of ion trap.

trode, is located at an angle to the optical axis of the tube. Acted upon by the transverse magnetic field in the region of bending of the electron gun, the electron beam is sharply deflected from its initial direction and is directed along the axis of the tube, while the stream of negative ions strikes the anode. Other ion-trap designs operate in similar fashion.

As we have already said, it is possible to protect the screen against damage from ion bombardment by aluminum treatment. With a sufficiently thick aluminum film, there is no need to utilize ion traps.

In addition to kinescopes for direct viewing, so-called projection picture tubes are also used; they are used in television receivers that have optical systems projecting the image from the screen of the kinescope onto large, flat screens.

Projection kinescopes (Fig. 15-32) have relatively small screen dimensions, and the imagery produced is magnified by an optical system

61-107

when it is projected on the flat screen. For adequate brightness of the picture to be viewed, screen brightness for the kinescope itself should be hundreds of times greater than that of direct-viewing kinescopes. Thus, projection picture tubes operate at considerably higher

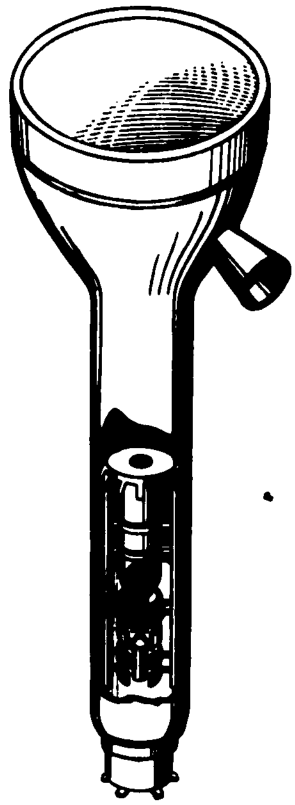

Fig. 15-32. Projection kinescope.

accelerating voltages, having values on the order of 20-50 kv. Operation at such high voltages dictates certain structural differences in projection kinescopes. Magnetic beam deflection is always used in such tubes, since with electrostatic deflection, such fast electrons require very high deflecting voltages. Focusing may be either electrostatic or mixed (with a second magnetic lens). The small screen dimensions of the projection kinescope necessitate narrow focusing of the beam, which is easily accomplished with magnetic focusing.

The modulator-diaphragm aperture is smaller than in standard

cathode-ray tubes which, first, decreases the size of the cathode emitting surface and, consequently, the diameter of the bright spot and, second, decreases the "penetrance" of the modulator for the field created by the high anode potential. The region between the anode and modulator is surrounded by a grounded cylinder, which guards against arcing over between anode and modulator. Voltage is supplied to the anode through a lead sealed into the conical section of the envelope, and connected to the conducting coating. In order to decrease leakage currents along the outer surface of the envelope, the anode lead is protected by a glass tube sealed into the envelope.

The screens of projection picture tubes, which give a white light, are always treated with aluminum, which protects them against rapid damage as a result of electron bombardment.

In using projection kinescopes, it is necessary to take into account the possibility of the appearance of X rays, created by the rapid deceleration of the beam electrons at the screen (bremsstrahlung X rays), and appropriate measures must be taken to guard against such radiation.

In recent years, work has been carried out on conversion to color television. The development of a system for reproducing a color picture is based upon the three-component theory of color vision, which states essentially that a visual perception of color for a light beam with any spectral composition may be reproduced with the aid of three single-color (monochromatic) light sources. By changing the ratio between the intensities of the component monochromatic beams, it is possible to obtain any colored visual impression.

Several different types of television picture tubes have been developed using this principle; of these, we shall discuss only one, namely the color kinescope using a shadow mask. In such a kinescope, excitation of each of the three primary colors is carried out by a

separate electron beam. The screen of the tube consists of regularly alternating red, green and blue dots of phosphor. Between the screen and the three electron guns creating the three independent electron beams, is a mask with apertures which prevents the electron beams from falling upon "foreign" color dots. The axes of the three electron beams should intersect in an aperture of the mask and, going on to diverge by a certain angle, the beam should fall on the appropriate phosphor dots (Fig. 15-33).

The screen of a color picture tube takes the form of a mosaic of

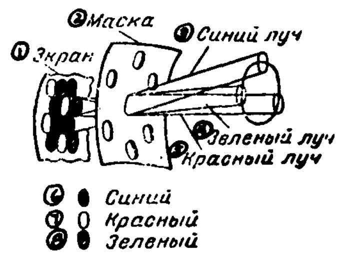

Fig. 15-33. Operating principle of color kinescope with aperture mask. 1) Screen; 2) mask; 3) blue beam; 4) green beam; 5) red beam; 6) blue; 7) red; 8) green.

color groups of phosphors, applied to the inside surface of the glass of the end of the tube. Each group consists of dots of red, green, and blue phosphors, located at the vertices of an equilateral triangle. In order to provide adequate resolving power for the tube, the number of such three-color elements must be quite large — about 400,000 — so that each dot will have a diameter of 0.4 mm. The screen and shadow mask have concentric convex surfaces with large radii of curvature.

The shadow mask takes the form of a convex metal sheet with circular apertures located opposite the centers of the color triangles on the screen. The number of apertures in the mask equals the number of color triangles on the screen.

For proper color reproduction, it is necessary to maintain rigid

mutual orientation of the electron guns, screen, and shadow mask, which is the basic complication in such a tube.

15-8. ELECTROOPTICAL IMAGE CONVERTERS.

Electrooptical image converters is the name given to instruments used to convert an optical light image of an object into an electronic image and then back to a light image. The initial light image of the object may be visible or invisible, depending upon the wavelength of the beams with which the object is illuminated, or which it produces (visible and invisible infrared or ultraviolet rays). Electrooptical converters, however, have found their greatest practical use in the conversion of images in infrared light.

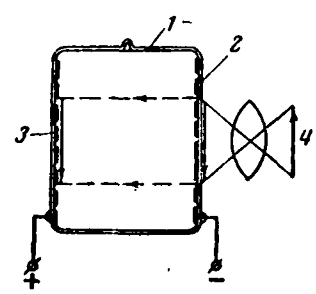

Fig. 15-34. Very simple converter design.

Let us examine the operating principle of the electrooptical converter by considering the very simple design shown in Fig. 15-34. On the inner surfaces of the ends of the cylindrical glass envelope 1 there is an applied semitransparent photocathode 2 that is sensitive, for example, to infrared radiation, and a luminescent screen 3, in turn coated with a thin metal layer (for example, a layer of aluminum). This metal layer should be transparent for photoelectrons, but opaque for light: light incident on the metal layer from the side of the observer (including screen glow), should be reflected from the metal layer. As a result, there is almost no illumination of the photocathode from the

direction of the screen.

A high direct-current voltage is applied between the photocathode and the screen. Infrared rays reflected from the object 4 are focused on the photocathode. Under the action of these rays, the photocathode emits photoelectrons, which are accelerated by the high voltage, and bombard the screen, causing the luminescing material to glow. With this arrangement, more electrons will be liberated from sections of the photocathode that receive more intense illumination. These electrons fall on corresponding sections of the screen, and cause intensive excitation of the phosphor. Those sections of the photocathode that are less intensely illuminated will yield fewer electrons. These electrons, falling on corresponding sections of the screen, will cause less intense screen excitations. Since different numbers of electrons will leave different sections of the photocathode, different numbers of electrons will be incident upon different sections of the screen. As a result, the sections of the screen will glow with varying intensities, reproducing the distribution of intensity received at the photocathode as a result of the light reflected from the elements of the object. The distribution of intensity over the screen accurately reproduces an image of the object, i.e., a visible picture of an invisible object is obtained on the screen; here the color of the picture is determined by the phosphor. Thus, in the electrooptical converter design under consideration, the light image of the object is transferred by the electrons directly from the photocathode to the screen with no use of electrooptical focusing.

The presence of a radial velocity component of the electrons, occasioned by the fact that a portion of the electrons leave the photocathode at various angles, leads to scattering of the electrons in the absence of an electrooptical focusing system. Each point on the photo-

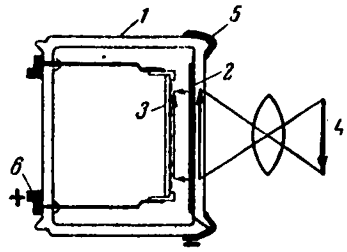

Fig. 15-35. Arrangement of improved converter.

cathode appears on the screen as a scattering circle, owing to which there is a deterioration in picture sharpness. The diameter of the scattering circle in this case is determined from the formula

$$d = 4l\sqrt{\frac{U_r}{U}},$$

where l and U are the separation and potential difference between photocathode and screen; U_r is the potential difference corresponding to the radial velocity component of the electrons.

It follows from this that some decrease in the scattering circle may be obtained by decreasing the screen-photocathode spacing, and increasing the voltage between them. With very small spacings, however, there may appear cold emission (autoelectron emission) from the photocathode, owing to which there will be a bright background on the screen. In addition, where l is very small, it is difficult to manufacture a sensitive photocathode.

An improved electrooptical converter is shown in Fig. 15-35. In the cylindrical glass envelope 1, which has two flat windows, there is located the semitransparent oxygen-cesium photocathode 2 on the screen 3 which takes the form of a thin glass disc on one side of which is applied a phosphor and a platinum grid that serves to increase screen conductivity. The distance between the photocathode and the screen

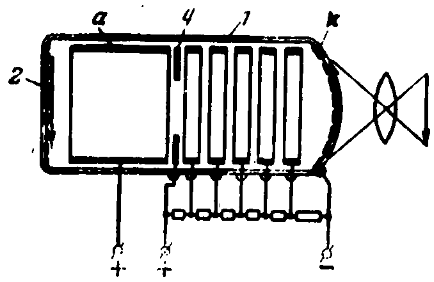

Fig. 15-36. Diagramatic representation of converter with electrostatic focusing.

equals about 5 mm. Since the path between the photocathode and the screen along the inside wall of the envelope is increased owing to the special shape of the envelope, in this design, there is good insulation between photocathode and screen, which makes it possible to utilize working voltages on the order of 5-7 kv. A working voltage is applied to contact rings 5 and 6. In order to increase insulation resistance, the inside surface of the envelopes is carefully cleaned.

In the converter design described, it is possible to utilize picture focusing, using for this purpose an electrostatic or uniform magnetic field set up by a long focusing coil. The use of focusing makes it possible to increase the separation of photocathode and screen, which permits a decrease in the bright background on the screen (where a non-metal-treated luminescent screen is used), while the use of a coil permits a decrease in the diameter of the scattering circle for the electrons. As a result, the quality of the image is considerably improved. Of course, the length of the focusing coil makes it impossible to change the size of the picture and, consequently, the region of application of the converter is limited. The use of a short focusing coil, which would make it possible to change the size of the picture, is not always possible since both long and short focusing coils require, in general, large-size converters and consume large amounts of power.

Thus, in practical application, we use the so-called electrooptical converters with electrostatic image focusing, which are smaller, draw slight amounts of power, and make it possible to change image size.

Converters with electrostatic focusing. The arrangement of an electrooptical converter with electrostatic focusing is shown in Fig. 15-36. In a cylindrical glass envelope 1 are located the screen 2 and photocathode k, which in order to improve picture quality (to decrease image distortion) is made somewhat convex, rather than flat, since minimum image distortion occurs with a convex photocathode, where the radius of curvature of the cathode equals the diameter of the cylinders of the electrostatic lens. The projection of the optical image of the object of the curved surface of the photocathode involves serious difficulties, however.

Near the photocathode there are several circular focusing electrodes, also designed to decrease image distortion. To these electrodes is applied a gradually increasing voltage from a potentiometer, which is normally located within the instrument. Between the ring-shaped electrodes and the screen there is an anode a, maintained at a high positive potential and playing the part of an accelerating electrode. Focusing is carried out by changing the ratio of the anode potential to the potential of the ring electrode nearest to it.

A change in picture size with this design of converter is made by changing the voltage on diaphragm 4, located between the last ring electrode and the anode. This voltage ranges from U_p (voltage on the last ring electrode) to U_a (anode voltage).

The utilization of various electrostatic focusing systems in converters makes it possible to improve quality or to simplify control of the converter.

Applications of electrooptical converters. Electrooptical con-

verters are utilized in various fields of science and technology for converting images of objects invisible in the dark into visible images. If the objects under study do not produce infrared radiation themselves, they are illuminated with such radiation, using standard incandescent bulbs, specially equipped and provided with infrared filters.

Especially wide use has been made of converters for various optical investigations such as, for example, investigation of the infrared radiation in the night sky, for determining the index of refraction for various materials in the infrared region, for investigating regions of the spectrum lying beyond the limits of human-eye visibility, etc.

Some types of converters are utilized to increase image brightness, in particular to increase the brightness of images on X-ray screens.

Electrooptical converters are utilized also in special optical systems that make it possible to magnify (bring nearer) an image of an invisible object. Such optical systems have arbitrarily received the name of electronic telescopes or binoculars.

15-9. TELEVISION TRANSMITTING TUBES.

Transmitting tubes operating in television studios are complicated devices representing a combination of photoelectric and cathode-ray

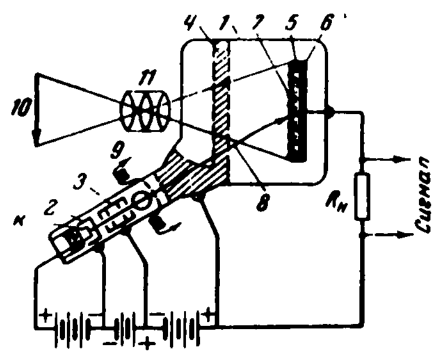

Fig. 15-37. Arrangement of iconoscope and basic external circuit.
1) Signal.

devices.

The iconoscope. A tube of this type utilizes the so-called principle of charge storage. The arrangement and circuit for the iconoscope is shown diagramatically in Fig. 15-37.

In the narrow tube of the glass envelope 1 is contained the electron gun, consisting of the cathode k, the control electrode 2, and accelerating electrode 3. On the inside surface of the envelope there is a metal conducting layer 4 that plays the part of a second anode and a collector of scattered electrons. In the wide portion of the envelope there is a composite photocathode, also called a mosaic or a mosaic screen. The mosaic consists of a thin mica plate 5 measuring about 100 cm^2 and 25-75 microns thick. On one side of this plate, there is a metal layer or thin metal plate 6 called the signal plate, since the videosignal is taken from it. On the other side of the plate there are several million globules of silver 7, insulated from each other, sensitized with cesium. These grains, also called elements of the mosaic, are very small globules of various diameters (5 microns or less). Each of these elements is on the one hand a miniature photocell and on the other a miniature capacitor, one plate of which, common to all the elements, is the signal plate, and the other is formed by the element itself. Thus, the entire mosaic (Fig. 15-38) may be considered to be a system of miniature capacitors, and at the same time a system of miniature photocells.

The electron beam 8 formed by the electron gun is incident upon the mosaic at a certain angle (Fig. 15-37). With the aid of the deflection coils 9, the beam is swept over the mosaic in lines, similar to the manner in which the beam is swept over the screen of a kinescope.

The beam travels over all lines, i.e., over the entire mosaic screen, at uniform speed in a period of 1/25 sec (time for a single

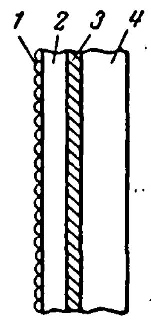

Fig. 15-38. Mosaic of iconoscope.
1) Mosaic elements (sensitized grains of silver); 2) thin mica plate; 3) metal backing layer; 4) massive support plate.

frame), while the beam flyback occupies a shorter period of time; during beam flyback, the beam is cut off by a negative voltage on the control electrode and, consequently, does not act on the mosaic.

The optical image of the object 10 is focused on the mosaic with the aid of the objective 11. Under the action of the light, the elements of the mosaic continuously emit photoelectrons whose number is proportional to the intensity of the light incident upon them. In accordance with the optical image, different elements of the mosaic receive different amounts of light and, consequently, emit different quantities of electrons. These electrons strike a positively charged collector 4, while each element of the mosaic, owing to the continuous emission of electrons, becomes positively charged. The more electrons emitted by an element, the higher the positive charge stored on it. Consequently, if the optical image of the object is focused upon the mosaic, a positive charge will be stored upon each element of the mosaic. The magnitudes of these charges are proportional to the intensity of the light

incident upon the elements, while their distribution over the mosaic surface is strictly in accordance with the distribution of light over the element owing to the projected image (image potential map).

If the mosaic is now scanned by an electron beam, the beam, as it passes over the lines, neutralizes the positive charges on the mosaic elements by means of its electrons. But neutralization of the positive charges on the elements liberates corresponding charges on the signal plate. As a result, a current appears in the signal-plate circuit whose magnitude varies as the beam moves over the mosaic elements, and is proportional to the change in the light incident upon these elements and, consequently, is proportional to the magnitude of the charges on these elements. When this current passes into the signal-plate circuit, a voltage drop appears across the load resistance R_n. Since discharging of the element capacitors by the beam occurs within a period of several microseconds, as the beam moves over the surface of the mosaic, current pulses appear in the signal-plate circuit, while across the load resistance there appear voltage pulses — the videosignals. Thus, the sequential illumination of the elements by the image is converted into a series of videosignals.

The efficiency of an iconoscope amounts all-in-all to about 10 per cent of its theoretical value. This low efficiency is explained by the fact that under the influence of the electrons in the scanning beam, which discharge the mosaic elements, secondary-electron emission occurs from the mosaic. The secondary electrons, leaving the mosaic at low velocities, return to neighboring elements, and diminish their positive charges that formed on them owing to photoelectron emission and, consequently, decrease the magnitude of the voltage pulses across the resistor R.

Secondary-electron emission decreases not only the efficiency of

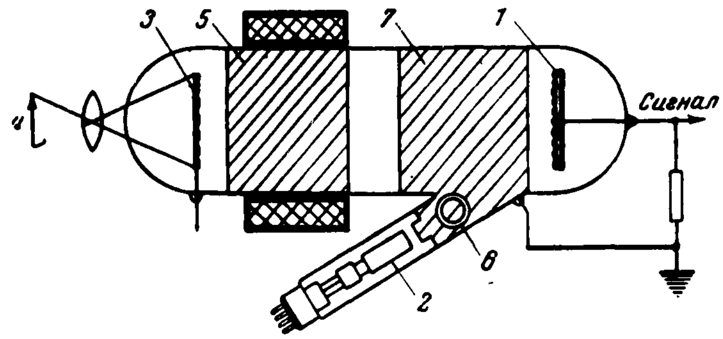

Fig. 15-39. Image iconoscope.
1) Signal.

the iconoscope, but its sensitivity as well.

In addition, in an iconoscope, photoelectric emission from mosaic elements is noticeably limited by the space charge. The presence of charge leakages along the surface of the mosaic also decreases efficiency and sensitivity in an iconoscope. Thus, iconoscopes are utilized only where the subjects to be transmitted are very strongly illuminated.

A substantial drawback of the iconoscope is the noticeable nonlinearity of the way in which sensitivity depends upon illumination.

The image iconoscope. In order to increase the sensitivity and efficiency of iconoscopes, various improvements have been made in them. As a result, a transmitting television tube has been created that is more complicated in design, and is called the image iconoscope. Since the image iconoscope is essentially a combination of a standard iconoscope and an electrooptical converter, it is also frequently called an image-transfer iconscope (Fig. 15-39).

Within the glass envelope are located the mosaic screen 1, electron gun 2, and semitransparent antimony-cesium (or oxygen-cesium) photocathode 3 upon which is focused the optical image of the subject 4. On the interior surface of the envelope there is a metal layer 5 which serves as the anode for the photocathode, and a metal layer 7 which is the collector designed to catch scattered electrons. The electron

beam is controlled with the aid of the deflecting coils 6.

Under light action from the focused image of the subject, photocathode 3 emits primary electrons, which are in turn focused, with the aid of the complex electron lens, on the mosaic screen 1. The electronic image focused on the mosaic causes secondary-electron emission from the mosaic elements, resulting in the appearance of positive charges on the elements. The secondary electrons leaving the mosaic of an image iconoscope as a result of the action of the electronic image play the same role as do the photoelectrons leaving the mosaic of a standard iconoscope as a result of the action of the optical image.

Since in contrast to the mosaic of an iconoscope, the mosaic of an image iconoscope becomes charged under the action of an electron image, rather than under the action of an optical image, it may be light-insensitive.

In the remaining details, the operating mechanism of the image iconoscope is similar to that of the iconoscope: an electron beam scans the mosaic, and the videosignals appear in the signal-plate circuit, and are then amplified and transmitted by radio.

The image iconoscope has many advantages in comparison with the iconoscope; the chief of these are: a) the secondary-electron emission current from the mosaic elements in the image iconoscope exceeds the photoelectric current from the mosaic elements of the ordinary iconoscope, and as a result the mosaic elements in the image iconoscope are charged to a higher potential than are the mosaic elements in an iconoscope, and thus the sensitivity of the image iconoscope may exceed the sensitivity of the iconoscope; b) the design of the image iconoscope makes it possible to utilize more sensitive photocathodes which also permits an increase in image-iconoscope sensitivity.

The image iconoscope possesses, however, one major disadvantage

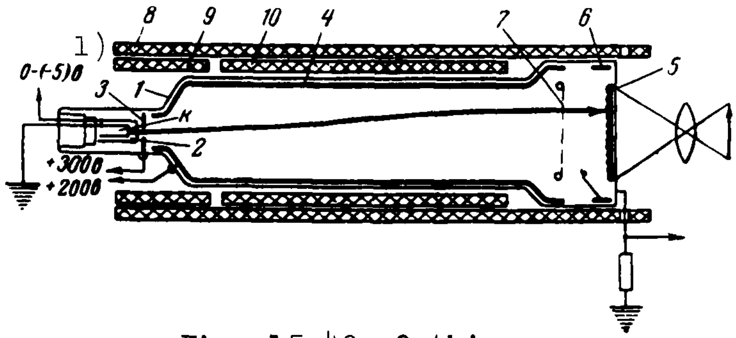

Fig. 15-40. Orthicon.
1) Volts.

which substantially limits its field of application. This consists in the fact that the electronic image in an image iconoscope, upon being transferred from the photocathode to the mosaic is noticeably distorted; distortionless transfer of the electronic image can occur only from a small central area of the photocathode.

The orthicon. In common with the iconoscope and the image iconoscope, the orthicon utilizes the principle of charge storage, but the design of an orthicon differs substantially from that of an iconoscope or an image iconoscope.

The basic arrangement of one orthicon design is shown in Fig. 15-40. In the narrow portion of the cylindrical envelope 1 is located the basic section of the electron gun, consisting of a cathode k, which is grounded, a modulator 2, to which is applied a potential varying from 0 to -5 v, the first anode 3, to which is applied approximately +300 v, and a second anode 4, whose potential equals about +200 v. The mosaic screen 5 is located on the inside surface of the flat polished end of the envelope, through which the connection with the signal plate is made. Near the mosaic screen on the inside surface of the cylindrical portion of the envelope there is a circular conducting coating 6, which is at ground potential, and which functions as a retarding electrode.

In front of the mosaic screen there is a fine metal grid 7, whose potential equals the second-anode potential (+200 v). This grid functions as a collector for the photoelectrons produced by the action of light on the elements of the mosaic, and for those electrons of the scanning beam that are slowed down near the mosaic and return to the grid.

The function of the coils outside the tube envelope are: focusing 8, alignment 9, and deflecting 10.

In the orthicon, the image of the subject is projected onto the mosaic from the flat end of the envelope, i.e., through the end of the envelope, the semitransparent metal layer of the signal plate, an insulator, and the thickness of the mosaic elements. This method of projecting the image of the subject upon the mosaic involves heavy absorption of light, up to 50 per cent, and the light is basically absorbed by the semitransparent signal-plate layer and within the mosaic elements. In contrast with the iconoscope and the image iconoscope, in the orthicon, scanning of the mosaic is carried out by a beam of slow electrons, which decreases the secondary-emission coefficient to 0.7.

There are two substantial drawbacks to slow-electron scanning. First of all, the slow electrons are subject to mutual repulsion, which leads to beam defocusing, and in addition, slow electrons are liable to be deflected toward those mosaic elements that have the greatest positive charges. In order to eliminate these defects, the orthicon utilizes a uniform longitudinal magnetic field, set up by a focusing coil.

Another feature of the orthicon is the fact that at a given light incidence on the mosaic, the photocurrent, which is determined by charge storage, reaches a saturation value. Clearly, at saturation current, large charges form on the mosaic elements, and when they are subsequently discharged by the electron beam to zero potential, considerable

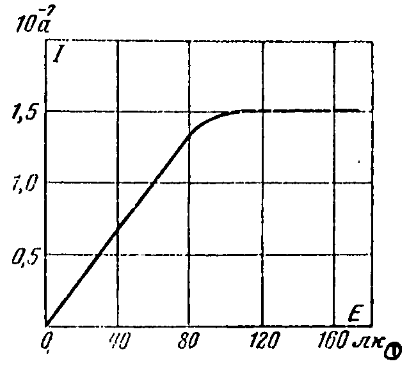

Fig. 15-41. Signal magnitude as a function of illuminance (amplitude characteristic) for an orthicon.
1) Lux.

current pulses (videosignals) will appear in the signal-plate circuit.

In the orthicon, scattered electrons are captured by a collector (grid), which is at a high positive potential with respect to the photocathode.

The orthicon is a superior transmitting cathode-ray tube in comparison with the image kinescope. The name of the tube is derived from the Greek word "ortho," which means "straight." This name underscores the linearity of the so-called amplitude characteristic of the tube, which represents the way in which the magnitude of signal current depends upon the magnitude of the light-flux or illuminance that is incident (Fig. 15-41). This linearity is observed with small values of illuminance. With large illuminance values, saturation occurs, since the current of the electron beam is inadequate to neutralize the charges appearing under the influence of the light on the mosaic elements.

The image orthicon. The television tube that utilizes the properties of the orthicon and the image iconoscope has come to be called the image orthicon (Fig. 15-42). The name "image orthicon" is also applied to the transmitting tube that uses a double-sided target, since its basic element is a thin glass plate (film), one side of

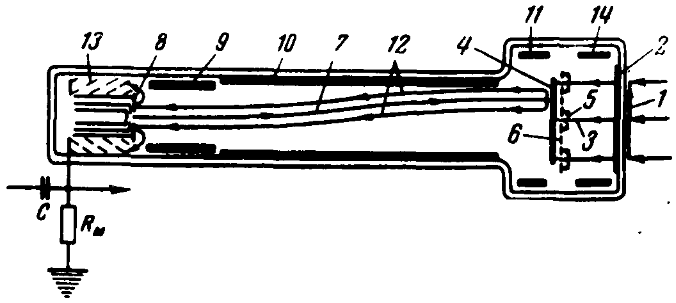

Fig. 15-42. Image orthicon.

which is subjected to the action of electrons emitted by photocathode, and the other to the action of the electrons in the scanning beam. This double-sided target functions as a mosaic screen.

In the image orthicon, the light image 1 from the subject is focused upon the semitransparent photocathode 2, which under the action of the light emits photoelectrons 3. These photoelectrons fall into a strong accelerating field, since there is a negative potential of -300 v with respect to the photoelectron collector applied to the photocathode. With the aid of an electron lens, the photoelectrons are focused on the double-sided target 4, which is a semiconducting glass film 0.005 mm thick.

As they strike the glass target at high velocities, the photoelectrons drive secondary electrons 5 out from its surface; they are drawn off by the grid 6, located at a distance of 0.05 mm from the target, and maintained at a potential of +1 v.

As a result of the departure of the secondary electrons, various sections of the target surface become positively charged. Thus, on the target surface, a potential map is formed that corresponds to the optical image of the subject on the photocathode.

Let us consider the mechanism by which the potential map on the target of an image orthicon is converted into videosignals. Each of the sections (elements) of the target surface, facing the grid, forms two

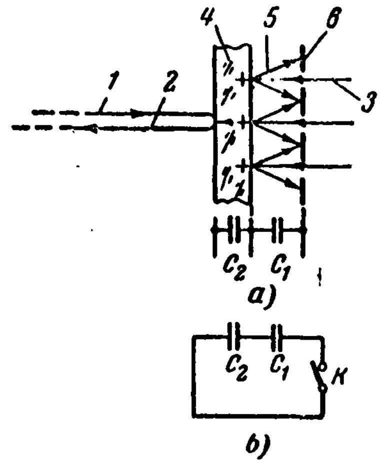

Fig. 15-43. a) Diagram of basic unit of superorthicon: scanning beam – target – grid; b) equivalent electrical circuit.

capacitors (Fig. 15-43a): a capacitance C_1 with respect to the grid, and a capacitance C_2 with respect to the section (elements) of the target surface facing the electron gun. Since the distance between target and grid is about ten times the target thickness, capacitance C_2 may be assumed to be considerably larger than capacitance C_1. The thickness of the glass target and its lateral resistivity are so chosen that in the time it takes to sweep out the image, i.e., the time for a single frame, the positive charges that form as a result of the departure of the secondary electrons gradually accumulate on the target elements. Under these conditions, capacitor C_1 is charged, while capacitor C_2 remains practically uncharged. When the elements of the rear side of the target are swept by the beam 1 (Fig. 15-43a), a redistribution of charges between capacitors C_1 and C_2 occurs. As a result, nearly all of the charge of capacitor C_1 is transferred to capacitor C_2. With a proper choice of thickness and resistivity of the target, the charge flowing off capacitor C_2 during the time between two successive scannings, should equal the charge stored during the same time on capacitor C_1. Thus, at the

moment of alternate switching, capacitor C_2 turns out to be free and, consequently, ready to receive the next charge. This pattern is repeated during successive cycles.

The arrangement of the basic image-orthicon unit that we have considered may be reduced to an equivalent electrical circuit, consisting of two series-connected capacitors C_1 and C_2, and a switch K (Fig. 15-43b), in which the switch performs the same function as the scanning beam in the image orthicon. If capacitor C_1 is gradually charged from any source, at the instant the circuit is closed, the charge on capacitor C_1 will be transferred almost completely to capacitor C_2, since C_2 is much greater than C_1. When the circuit is opened, capacitor C_2 discharges (owing to compensation of its charge by the electrons of the scanning beam during the switching instant), and capacitor C_1 is charged (owing to the source of power). With circuit closings repeating at specific time intervals, which are equal to the intervals of time between successive cycles in the image orthicon, the picture will be repeated.

Thus, before switching, the scanned side of the target contains no charges, while at the instant of switching, there are stored on it the charges of which a portion interact with the electrons of the scanning beam, and the remaining portion flow off along the target.

In the image orthicon, sweeping of the target is performed by slow electrons, created as follows: The electrons leaving the cathode of the electron gun are accelerated by the fields of the first 8 and second 9 anodes, whose potentials equal $\pm$210 v, but are retarded in the region of the third anode 10, whose potential equals +125 v (Fig. 15-42). As they near the target, these electrons are additionally slowed down by the fields of the retarding electrode 11, at a potential of +25 v, and the target, whose potential equals zero, approximately. Thus, a beam of slow electrons 7 is formed, which is used to scan the target.

The velocity of these electrons is so low that the secondary-electron emission coefficient, as in the orthicon, does not exceed 0.7-0.8.

As they fall upon an element of the target, the electrons of the scanning beam neutralize its positive charge. Depending upon the magnitude of the charge upon it, more or fewer electrons of the scanning beam will be required for neutralization.

If the charge on any element is slight, a portion of the electrons from the scanning beam is expended in neutralizing the charge, while the remainder (excess) electrons are returned back to the electron gun, and form a back electron beam 12 (Fig. 15-42) or 2 (Fig. 15-43a). Since the charges on the various target elements differ, different numbers of electrons will be required for the scanning beam to neutralize them and, consequently, different numbers of electrons will be returned from different elements as they form the return beam. Thus, the return-beam current turns out to be amplitude-modulated in accordance with the distribution of the charges over the target elements: large charges correspond to low return-beam current, and conversely.

Since maximum illumination of a point on the photocathode corresponds to a target element with maximum positive charge, and in neutralizing such a charge, a larger number of electrons is required from the scanning beam, the number of electrons in the return beam will be smaller. Thus, the brightest point on the object being transmitted will correspond to minimum return-beam current.

In moving from the target toward the cathode of the electron gun in an accelerating field, the return-beam electrons, which form the videosignal, attain high speeds. As they impinge upon the anode of the electron gun, which is at the same time the first emitter in a secondary-electron multiplier 13 (located around the first anode of the gun), the electrons of the return beam drive out secondary electrons, which in

turn strike the second emitter, and drive out from it a still greater number of secondary electrons, etc. By the time it is taken off from the anode of the multiplier, the videosignals have already been amplified by about a factor of 500.

The use of the secondary-electron multiplier permits a considerable increase in the sensitivity of an image orthicon. In tubes without such a multiplier, the sensitivity is limited by the noise appearing in the tube itself and in the electron tubes of the videosignal amplifier. Where there is an electron multiplier within the transmitting tube itself, the sensitivity is essentially limited only by the noise in the tube itself.

The image orthicon has about 100 times the sensitivity of an orthicon. Thus, the image orthicon may operate under natural lighting, and requires no additional light sources.

The image orthicon has several substantial faults, of which the basic is its high noise level. As the illumination of the object drops, i.e., as the input signal decreases, the noise level rises. This is explained by the fact that the darkest point on the object image corresponds to maximum output current, and, consequently, to maximum noise level since as the current increases, noise rises.

The complicated manufacturing processes for an image orthicon is also a drawback.

The vidicon. The vidicon is a transmitting television tube in which an internal photoeffect is used in place of external photoemission.

The basic part of a vidicon (Fig. 15-44) is the photoresistor 1, which is applied to a semitransparent metal layer 2, which is the signal plate. The photoresistor has a sensitivity of 300 μa/lumen, i.e., it is ten times more sensitive than photocells using external photoemission.

In the vidicon, the optical image 3 of the subject is projected through the input glass of the tube 4 and the semitransparent metal

layer 2, which is the signal plate, onto the photoresistor 1. The other side of the photoresistor is scanned by electron beam 5, which is focused by a uniform magnetic field set up with the aid of focusing coil 6. In addition to the focusing coil there are alignment 7 and focusing 8 coils around the tube envelope.

The semitransparent metal layer of the signal plate is maintained at a constant potential with respect to the cathode of the electron gun. This potential is set at from 10-30v. In front of the photoresistor there is located a grid 9 that is transparent to electrons, and that sets up a uniform electric field along the section between the photoresistor and the grid.

If no light falls on the photoresistor, its resistivity is no lower than 10^{12} ohm·cm. Thus, only a small dark current flows through the photoresistance. If a light-image of an object to be transmitted is

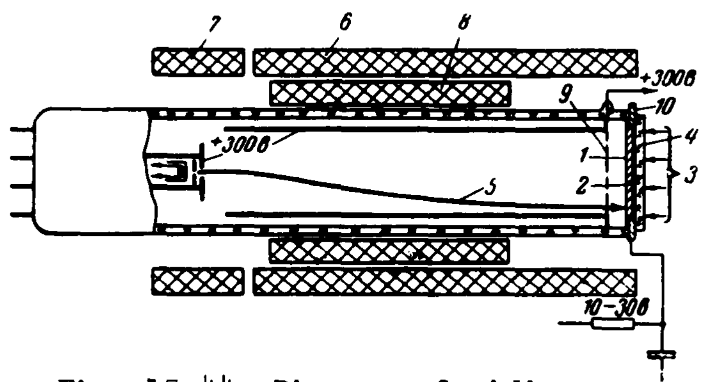

Fig. 15-44. Diagram of vidicon.
1) Volts.

projected on the photoresistor, its conductance rises sharply, but it will be different in different sections, depending upon the illumination. As a result, various sections of the scanned surface of the photoresistor are charged with respect to the cathode to various positive potentials, reaching +2 v. The values of these potentials depend upon the time required to scan a single frame: the greater the frame-scanning time, the higher the potential, i.e., the higher the positive charge

stored on one element or another of the photoresistance.

Upon neutralization of the positive charges on the photoresistance elements by the electron beam, a negative videosignal current pulse appears in the signal-plate circuit.

In the poor lighting in which a vidicon is normally used, its light characteristic (dependence of output signal upon light incident upon the tube) is linear, but for large values of illumination, saturation takes place.

The arrangement and adjustment of the vidicon is simple. The high output-signal level does not require amplification by means of electron multiplier. There is no need in the vidicon for electrooptical image conversion, either, since the inherent sensitivity of this tube is quite high. Under normal operation, a current on the order of 0.2 μa should flow through the vidicon photoresistor; this provides a signal-to-noise ratio of roughly 100, i.e., the vidicon gives noiseless operation. The enumerated advantages of the vidicon make it possible to transmit with illuminances on the order of several luxes, i.e., under normal industrial illumination conditions. The dimensions of the vidicon are comparatively small (diameter 25 mm, length 150 mm).

The basic drawback to the vidicon, which is a property of its photoresistance, is its slow response.

Chapter Sixteen

X-RAY TUBES

16-1. X RAYS: THEIR NATURE AND PROPERTIES.

In 1895, studying the properties of cathode rays (high speed electrons), Roentgen observed that new, invisible rays were emitted from those parts of a glass tube struck by cathode rays. Roentgen observed these rays by the luminescence of a screen consisting of a layer of crystalline particles of barium platinocyanide. The screen glowed when the invisible rays struck it. Roentgen named the new rays, whose properties and nature were as yet unknown, X rays. Later they were named roentgen rays in honor of Roentgen.

X rays appear at any time when the flow of highly accelerated electrons meets any body in its path. Especially strong X rays are obtained in those cases when the fast-moving electrons meet the heaviest metals, such as, for example, platinum or tungsten, in their paths.

The detailed study of X rays has established the great diversity of their properties. First of all, these invisible rays can penetrate many bodies which are opaque to visible light; they are absorbed and scattered in different ways by bodies which have different densities. X rays ionize gases, making them electrical conductors; affect photographic plates; cause fluorescence of some substances; and show a strong physiological effect on living organisms (they destroy living cells!). Finally, X rays do not carry any sort of charge with them; therefore, they are not deflected by either magnetic or electric fields, and, consequently, are always propagated in straight lines.

When X rays pass through different mediums their intensity is reduced according to the exponential law $I = I_0 e^{-\mu d}$, where I_0 is the

intensity of incident uniform (fixed wavelength) X-ray radiation, I is the intensity of X-ray radiation after passing through the layer of matter, d is the thickness of the layer, and μ is the coefficient of absorption of the given matter.

The described properties of X rays show that they resemble neither cathode nor anode (flow of positive ions) rays. At the same time, they scarcely resemble light rays, since they are not strictly subject to the laws of reflection and refraction established for visible light. However, thorough investigation of X rays has established that their nature is like that of visible light; the difference lies only in that X rays are electromagnetic oscillations with a very short wavelength (from several hundred A to 20 A) and, consequently, they have very great energy. The quantum energy of X-ray radiation is on the order of tens or hundreds of thousands greater than the quantum energy of visible light.

The wavelength of X rays depends on their excitation conditions. Soft X rays, with a wavelength of from several A up to 20 A, differ from hard rays, whose wavelengths amount to tenths or hundredths of A. Besides this, it has been established that the shorter the wavelengths, the harder the X rays and the deeper they penetrate into bodies.

Since the wavelength of X rays is considerably less than the wavelength of visible light, ordinary glass surfaces, which are smooth under visible light, are rough under X rays, just as a matted glass surface is rough under visible light. This is explained by the fact that the laws of reflection and refraction of light optics are inapplicable to X rays.

16-2. EXCITATION OF X RAYS.

X rays are excited when electrons which have been highly acceler-

ated by an electric field strike an anode. Two types of rays may be excited by this means: bremsstrahlung X rays and characteristic X rays.

a) Excitation of bremsstrahlung X rays.

If the electrons traverse a great potential difference of the accelerating electrical field U_a in the space between the cathode and the anode, at the end of the path they will have acquired great kinetic energy, $mv^2/2 = eU_a$. Since the initial kinetic energy of the electrons, which they had at exit from the cathode, is small, then, neglecting it, one may consider that all the electrons have the same energy, eU_a, at the end of the path. Striking the anode, these electrons meet free and bound electrons of metal in their path, and, reacting with them, are sharply decelerated. In deceleration, the kinetic energy of the electrons is converted into electromagnetic oscillation energy of very short wavelength; that is, into X-ray radiation energy. Such radiation has been named bremsstrahlung X-ray radiation, or simply bremsstrahlung X rays.

We will designate the energy of the quantum of X-ray radiation by $h\nu$. Then for complete conversion of the kinetic energy of the electrons into radiation energy we will obtain the equation

$$\frac{mv^2}{2} = eU_a = h\nu_{\text{макс}} = h\frac{c}{\lambda_{\text{мин}}},$$

from which we may determine the maximum energy or the minimum wavelength λ_{min} corresponding to it by:

$$\lambda_{\text{мин}} = \frac{hc}{eU_a} = \frac{12{,}35}{U_a},$$

where λ_{min} is the wavelength of X-ray radiation, in A, and U_a the anode voltage in kilovolts.

It follows from this expression that if the electrons have identical kinetic energies at the moment they strike the anode, the wave-

length of the bremsstrahlung X-ray radiation must be one and the same (monochromatic radiation); that is, the same spectral line as for the bremsstrahlung X-ray radiation must be obtained. In reality, in deceleration the kinetic energy of only some electrons is entirely converted into X-ray radiation energy. The energy of the other electrons is only partially converted into X-ray radiation energy; the remaining part of these electrons' energy is converted into heat energy, which heats up the anode. Therefore, actually, bremsstrahlung X rays do not consist of waves of one length, but of numbers of waves of different lengths ranging from λ_{min} to λ_{max}; radiation on wavelength λ_{min} is obtained by complete conversion of the kinetic energy of the electrons into the quanta of X-ray radiation. This minimum wavelength also determines the boundary of the spectrum of bremsstrahlung X-ray radiation.

Naturally, the intensity of X-ray radiation must be different for each value of wavelength, since it is determined by the number of quanta possessed of this or that energy, that is, this or that wavelength. The magnitude of the quantum energy is determined by the portion of the kinetic energy of the electron which entered into the formation of the quantum. Using a massive anode with the atomic number Z and constant values of anode current and voltage in the tube, the form of the distribution curves of bremsstrahlung X-ray radiation, in a spectrum, to good approximation, may be described by the equation

$$I_{\lambda} = kIUZ\frac{\lambda - \lambda_0}{\lambda^3} \; [\text{эрг/см}^2 \cdot \text{сек}].$$

Intensity distribution in the spectrum of bremsstrahlung X-ray radiation for different values of anode voltage shows that (Fig. 16-1), in the first place, intensity distribution of bremsstrahlung X-ray radiation corresponds to the continuous spectrum and, in the second place, the commencement of the curves (λ_0) and their maxima move

toward the left, that is, toward the shorter wavelengths, as anode

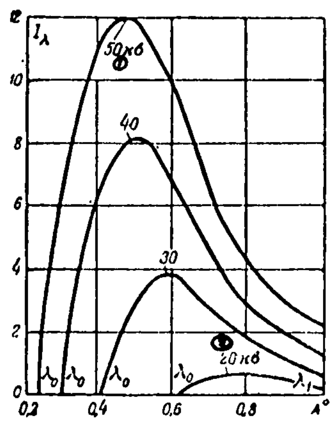

Fig. 16-1. Intensity distribution, in a continuous spectrum, of bremsstrahlung rays for a tungsten anode in the presence of different anode voltages and constant current. 1) 50 kv; 2) 20 kv.

voltage increases. This means that the greater the anode voltage, the harder and stronger the X-ray radiation becomes. In the presence of very high anode voltages, the electrons acquire enormous speeds and excite, together with bremsstrahlung X-ray radiation, the so-called characteristic X-ray radiation, as they penetrate atoms.

b) Excitation of characteristic X rays.

If an electron has great energy and meets an atomic electron shell in its path, it may be decelerated, by this means releasing part of its energy in the form of bremsstrahlung radiation energy and part of the energy in the form of heat. If the energy of the moving electrons is too great, then, decelerating and converting part of its kinetic energy into bremsstrahlung X-ray radiation energy, it penetrates into the atom and releases a considerable part of its remaining energy to one of the electrons of the atom, which, having received this energy, may leave the atom. The place of the electron which has left may be occupied by another electron of the same atom which is located on a higher energy level, or by a free electron of the anode metal, radiating

by this means a quite definite quanta of energy in the form of X rays or of light. Since electrons may be dislodged from various energy levels of the atom, then, when their places are occupied by other electrons, the atoms will emit energy corresponding to these electron transitions; that is, the energy of radiation on quite definite wavelengths characteristic for electron transitions in atoms of the given type.

Since the structure of the electron shells of atoms of different chemical elements is different, then every chemical element has a spectrum which is characteristic only of it. Therefore, the X-ray radiation obtained in a given case is called characteristic; it is not a continuous spectrum, as in the case of bremsstrahlung X-ray radiation, but only separate spectral lines.

We will consider the theory of origin of characteristic X-ray radiation, as formulated in 1916 by Kossel on the basis of the quantum theory of the structure of the boron atom, with the uranium atom as an example.

The electrons revolving around the nucleus of the uranium atom may be divided into separate groups, within which they have equal energies. These groups form the so-called electron layers (or shells), conventionally designated by letters K, L, M, N, O, P, and correspond to definite levels of energy; layer K, located closest to the nucleus, corresponds to one level of energy (the lowest energy level), layer L consists of three levels of energy (L_1, L_2, L_3), layer M, five levels (M_1 through M_5), layer N, seven (N_1 through N_7), layer O, five (O_1 through O_5). The most distant from the nucleus, layer P, consists of three levels of energy (P_1, P_2, P_3); these are the highest levels of energy. Thus, 24 energy levels are distinguished. The number of energy levels is determined by the amount of electrons in the atom and varies from one to 24. For example, the uranium atom is characterized by 24

levels of energy (Fig. 16-2), conventionally indicated in the text figure in the form of circles, which, of course, have nothing in common with the actual complex orbit of the electrons.

With the aid of the cited schematic of the energy levels of the atom it is easy to explain the origin of characteristic X-ray radiation. The fact is, if a free electron, which has a great enough energy, penetrates into the atom, it may displace an electron from any level (for example, from level K). If, by this means, a free electron and an electron which has been displaced from level K are removed from the atom, then a shell vacancy is created in layer K, and the atom becomes a positive ion. A metallic atom cannot remain in such a state for long. One of the electrons which belong to the higher energy level or a free electron from the periphery of the atom shifts to the free space in layer K. In such transition, as we know, the atom radiates energy, the quantity of which is equal to the difference between the energy of the electron located in the higher level, and its energy in the level to which it has shifted (for example, from level L to level K). By this means the atom emits one wave of radiant energy, that is, it gives one spectral line.

Since electrons may shift to level K from any of the other overlying levels, then the atom will radiate energy in the form of quanta with various wavelengths corresponding to the fixed electron shifts. The series of spectral lines corresponding to the wavelengths originated in electron transitions from the overlying energy levels to level K are called characteristic X-ray lines of the K series spectrum. Similarly, electron transitions from overlying levels to level L give spectral lines which are called L-series lines and so forth.

Characteristic X rays of the K-series are harder than rays of

other series. This is explained by the fact that in the transition, for example, of electrons from the periphery of the atom to the K-level the atom radiates the greatest energy, corresponding to the

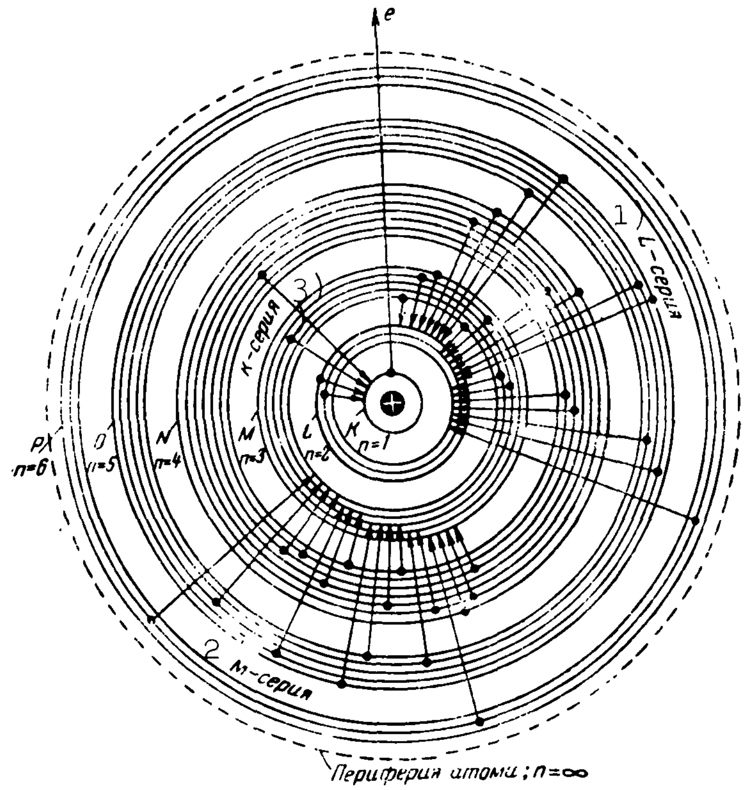

Fig. 16-2. Conventional diagram of energy level distribution and electron transitions accompanying irradiation (for the uranium atom). 1) L-series; 2) M-series; 3) K-series; 4) periphery of the atom; n = ∞.

quantum with the shortest wavelength.

Characteristic X rays are created in electron transitions from the higher energy levels to the lower, since in this case the difference between the energies corresponding to these levels is greatest, and the wavelengths of the radiated quanta of energy are small. If transitions take place from the initial levels to others which are little different in their energies from the initial levels, then the quanta of the radiant energy are small, and the wavelengths great.

In such transitions the atoms radiate lines in the visible part of the spectrum, that is, ordinary visible light.

In the heavy atoms the layers closest to the nucleus have identical structure. Electron transition in these layers must give a fixed hardness of radiation, that is, an equal amount and location of spectral lines of characteristic X-ray radiation. However, the hardness of characteristic X-ray radiation does not depend only on electron transition, but also on the atomic number of the element: the greater the atomic number of the element, the greater the charge of the nucleus and the stronger the electrical field around it. Therefore, the difference of energy between analagous layers in different atoms also increases in proportion to the atomic number of the element.

Characteristic X-ray radiation originates simultaneously with bremsstrahlung X-ray radiation in the presence of higher anode voltages. The quantity of the anode voltage for this must be not less than the so-called critical value $U_{a.k}$,* which is determined from the equation

$$U_{а.к} = \frac{12{,}35}{\lambda_к},$$

where λ_k is the least wavelength of the series of spectral lines, which are excited by the critical value of anode voltage; the wavelength of the spectral lines from the anode voltage does not depend on them but depends only on the electron transition in atoms of the substance in which the electrons are decelerated.

Practically speaking, in the presence of comparatively low anode voltages, it is possible to obtain only bremsstrahlung X-ray radiation in a pure form.

Knowing these spectral lines, we can say to which substance they

---

* [$U_{a.k} = U_{a.k} = U_{critical\ value}$.]

belong, that is, identify the substance by its spectral lines.

Characteristic X-ray radiation has considerably less energy than bremsstrahlung X-ray radiation. Practical use of X-ray radiation, in the overwhelming majority of cases, is made only of bremsstrahlung rays. Characteristic X rays are used only in X-ray defraction and X-ray spectral analyses.

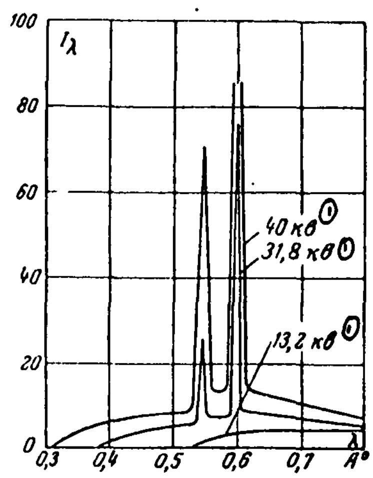

Fig. 16-3. Distribution of spectrum intensity of bremsstrahlung and characteristic radiation for a tube with rhodium anode.
1) kv.

The line spectrum of characteristic radiation and the continuous spectrum of bremsstrahlung radiation for a tube with a rhodium anode with anode voltages of 13.2, 31.8, and 40 kilovolts are shown in Fig. 16-3. The intensity distribution curves of the X-ray radiation show, according to the spectrum, that with a voltage of 13.2 kv only bremsstrahlung X-ray radiation, giving a continuous spectrum, is excited. With a voltage of 31.8 kv in the tube, together with bremsstrahlung radiation, giving a continuous spectrum, there is also excited characteristic X-ray radiation giving a line spectrum, and

superimposed on the continuous spectrum of the bremsstrahlung radiation. With a voltage of 40 kv, the intensity distribution of the radiation, according to the spectrum, remains qualitatively the same as at 31.8 kv. From this it follows that with the increase of anode voltage the intensity of bremsstrahlung and characteristic X-ray radiation is increased, but the wavelength of characteristic radiation remains invariable. Besides this, bremsstrahlung radiation becomes harder as anode voltage increases.

16-3. PRACTICAL USAGE OF X RAYS.

The properties of X rays have predestined their widespread application in science, technology, and medicine. X rays are used for diagnosis (X-ray examination) in medicine. By means of fluoroscopy the structure and arrangement of the internal organs of the patient are studied, and their defects are revealed. For example, by means of X-ray examination, diseases of the lungs and teeth are diagnosed and the gastrointestinal tract is inspected.

Placing, for example, a hand between a fluorescent screen and an X-ray tube, one may see the shadow image of the bones on the screen, since bones attenuate X rays more strongly than soft tissue. This image may be focused on a light-sensitive film and an X-ray photograph obtained, which facilitates the diagnosis of diseases of the bone, discovery of foreign objects in the human body, and so forth. Fluorescent screens which, under the action of X rays, luminesce with a green color are used in X-ray examination and visual observation, because the human eye is most sensitive to this color.

X rays of fixed hardness and intensity are used in medicine and for medical purposes (X-ray therapy). For example, with the aid of X rays, skin diseases and some types of cancer are treated; the healing of open wounds is expedited, and so forth.

For medical treatment, tubes are used in which a narrow beam of X rays is utilized. This beam is concentrated only on the affected parts of the human organism.

In technology, X rays are widely utilized for the inspection, for example, of the quality of casting and welding and for X-ray diffraction and X-ray spectral analyses.

X-ray inspection of materials or finished parts (X-ray defectoscopy is made with the aid of very hard X rays, which, penetrating to a great depth and, being absorbed differently by a substance with a differing density, permit the observation of possible heterogeneities and defects without destroying the inspection sample.

If X rays traverse through metal, inside of which there is, for example, a blister, the rays which do not meet blisters in their path and traverse the entire thickness of the metal are absorbed to a greater degree than the rays which have met a blister in their path. The attenuated rays operate more weakly on a fluorescent screen placed behind the metal than do the undiffused rays and the blister is reflected as a lighter patch of corresponding shape on the darker background of the remaining metal. Detailed visual X-ray examination is not sufficiently accurate. A more sensitive method of examination, by means of selection of exposure and tube anode voltage, is to obtain an X-ray photograph of the inspection sample.

Soft X rays are not suitable for X-ray examination since they are strongly absorbed by metal.

X-ray diffraction analysis, in which the scattering effect of X rays by crystalline substances is used, is carried out, as a rule, with the aid of monochromatic X rays, that is, rays of one and the same wavelength. The scattering is caused by the fact that the distance between adjacent atoms of the space lattice and wavelength of the X rays

have the same order; accordingly, the X rays may be in certain conditions reflected from the surfaces of the space lattice of the crystal. These conditions are the fixed relationships between the wavelength of the incident X rays, the angle of their incidence on the surface of the crystal, and the distance between the two adjacent reflecting surfaces.

In X-ray diffraction analysis, the X rays are aimed at the inspection sample, which has been placed in a special chamber. The X rays reflected at fixed angles from the various surfaces of the space lattice of the crystal of the sample strike a light-sensitive film, on which, after development, dark lines or patches, depending on the method of analysis, are observed. These lines (or patches) are arranged in accordance with the X rays reflected from the crystal and falling on the film. The structure of the inspection sample is determined from their arrangement.

In structural analysis either bremsstrahlung rays or characteristic rays may be used. In the latter case, the material of the anode of the X-ray tube must be such that the wavelength of the X-ray radiation is close to the distance between the two adjacent reflecting surfaces of the crystalline lattice of the substance being inspected. For this reason, anodes or tubes for structural analysis are made of various materials, and anode voltage is selected in accordance with the requirements of the analysis. Anode voltages in use vary from 15 to 100 kv.

Quantitative and qualitative analysis of the elements included in the material being studied is carried out by the X-ray spectral analysis method. In most cases, in X-ray spectral analysis the so-called emission method is used, in which the inspection sample, placed at the anode of the X-ray tube, is exposed to electron bombardment and the characteristic X-ray radiation produced by this means is studied with the aid of the X-ray spectrograph. In the spectrograph, the X ray,

in passing through the crystal, are scattered, and then, as in structural analysis, act upon a light-sensitive film.

Voltages and exposures used in the emission method are much the same as those used in structural analysis.

16-4. THE ELECTRONIC X-RAY TUBE: ITS CHARACTERISTICS AND PARAMETERS.

The electronic X-ray tube, in its simplest form, consists of a glass bulb, inside which two electrodes are located: the cathode (1) and the anode (2) (Fig. 16-4). The air is carefully evacuated from the bulb (the pressure of the remaining gases must not exceed 10^{-6}-10^{-7} millimeters of mercury. The cathode incandesces to working temperature and serves as the source of the electrons which are required for excitation of X-ray radiation at the anode. High positive voltage, usually tens or hundreds of kilovolts, is fed to the tube's anode to give the electrons great speeds. The higher this voltage, the greater the penetrating power the X rays possess. In tubes working at voltages below 500 kv, the anode is made solid and cut at an angle of 20 to 45° toward the direction of movement of the flow of electrons (Fig. 16-5). With such an anode cut, the X rays have the greatest intensity in a solid angle of about 25 to 30°, circumscribed around a normal toward the central part of the working surface of the anode. With increase of the solid angle the intensity of the X-ray radiation decreases (Fig. 16-6), since the X rays excited in the anode traverse greater distances in these directions in the body of the anode and are more strongly diffused than in the direction perpendicular to the anode. In a perpendicular direction the X rays traverse distances in the body of the anode which are equal to the depth of penetration of the electrons in the anode (approximately 10^{-3} mm).

To obtain strong X-ray radiation, the electrons departing from

the cathode are focused on the working plane of the anode. In its simplest form, such a focusing device is a cylinder, inside of which the cathode is placed. Sharpness of focusing depends on the location of

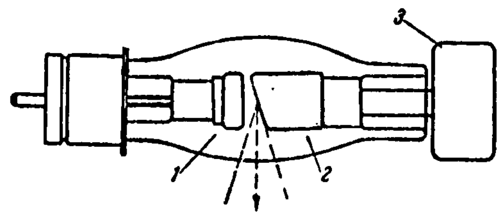

Fig. 16-4. Structure of an X-ray tube with radiation cooling.

the cathode in the focusing cylinder: sharper focusing is obtained by placing the cathode deeper in the cylinder.

The part of the working surface of the anode on which the focused beam of electrons falls is called the focus of the tube. The focus is circular or rectangular, determined by the shape of the cathode and the shape of the focusing device. The excited X rays are propagated from the tube focus in a straight line by the outgoing beam.

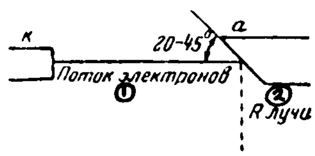

Fig. 16-5. Direction of electron flow and X-ray flow.
1) Flow of electrons; 2) X rays.

To increase the intensity of X-ray radiation, X-ray tubes are made with small strongly marked focuses which form a small focus spot of the required shape in the middle of the working surface of the anode.

The greatest density of electrons, whose kinetic energy in deceleration is converted into X-ray radiation and heat energy, is created in these focal spots. By this means, at best, several per cent of the entire

kinetic energy of the electrons changes, during their deceleration, into X-ray radiation energy, and the greatest part of the energy (of the order of 98 per cent) changes to heat energy. As a result, the

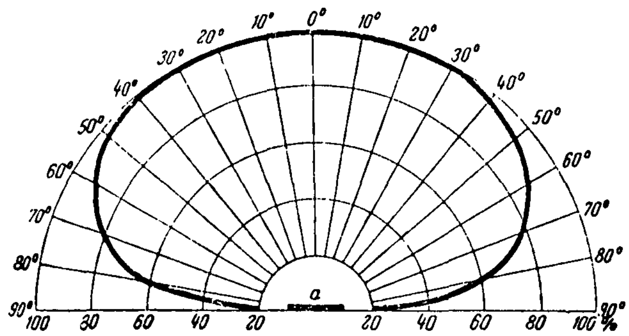

Fig. 16-6. Spatial distribution of X-ray radiation intensity.

anode heats up to a very high temperature and discharges a great quantity of gases. As a consequence of this, the operating conditions of the tube are disturbed. For normal tube operating conditions, by which the greatest possible intensity of X-ray radiation is obtained, this heat must be removed from the focal point. For this purpose, the anode body is made from copper and cooling is applied.

If the tube's construction does not provide for a cooling device, its anode is made of high-melting metal (usually of tungsten), which, when the tube operates, heats up, and scatters the heat into the surrounding space by means of radiation (natural cooling). In powerful tubes, anodes with a hollow stem are used. A metallic reservoir, whose cavity is connected with the cavity of the stem, is fastened to the outer end of the stem (Fig. 16-7). The reservoir is filled with distilled water. When the tube operates, the water in the hollow of the stem of the anode heats up and by its natural circulation, the heat is removed from the anode to the reservoir, and transferred to the surrounding space through its wall. More intensive cooling is attained by

forced circulation of water or oil, accomplished with the aid of a pump or running water from a pipe line. In the latter case the anode of the X-ray tube is grounded.

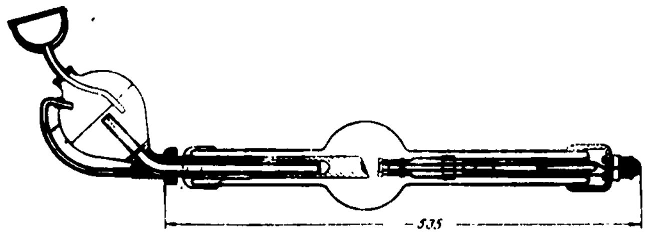

Fig. 16-7. 110-kv X-ray tube with water cooling for operation in air.

In low-power tubes radiation cooling is used; the stem of the anode is made continuous and a radiator (3 in Fig. 16-4) is fastened to it from the outside to remove heat from the anode to the surround-

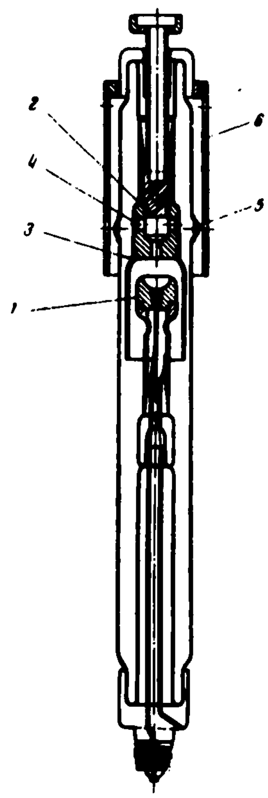

Fig. 16-8. X-ray tube for structural analysis.

61-10

ing space.

A significant rise of anode temperature limits the permissible load of the X-ray tube. Modern X-ray tubes have a continuous permissible load of not more than several kilowatts.

The primary electrons, focused and highly accelerated, moving from the cathode to the anode may cause not only X-ray radiation from the anode, but also a secondary electron emission. By this means part of the primary electrons may be reflected from the anode. Also, the reflected electrons, striking the internal surface of the glass bulb of the tube with great speed, heat it up and cause an additional secondary electron emission from it. Besides, the primary electrons, striking the internal surface of the bulb, form a negative surface charge on it. As a result, the tube's operating conditions deteriorate noticeably. To eliminate the harmful influence of the electrons striking the internal surface of the bulb, metal screens are used, by means of which both the primary electrons reflected from the anode and the secondary electrons discharged from the anode are caught. The typical construction of an X-ray tube with a protective screen (a copper sheet fastened to the anode which catches electrons moving from the anode to the bulb) is shown in Figs. 16-8 and 16-10. The rays pass through holes in the sheet, in which, in some cases, a thin special plate of beryllium is inserted, which arrests the primary electrons reflected from the anode, but has practically no absorption effect on X rays.

The intensity of X-ray radiation depends on anode voltage, anode current, and atomic number (Z) of the element of which the anode is made. The character of this relationship is determined by the formula $I_\lambda = kU_a^2 I_a Z$, where $\underline{k}$ is the factor of proportionality depending on the construction of the tube. Therefore, to increase the intensity of X-ray radiation it is necessary to increase not only the anode voltage and

anode current, but also the atomic number of the anode's element, that is, it is necessary to make the anode of high-melting metal with large atomic numbers, for example, tungsten or platinum. However, practically

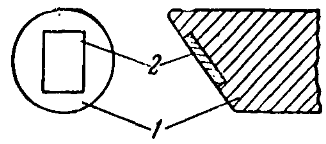

Fig. 16-9. Anode of an X-ray tube.
1) Anode body; 2) anode mirror.

speaking, in many cases it has been found more advantageous to make the anode of copper, and solder into its body the so-called mirror, a plate of the corresponding heavy metal, as shown in Fig. 16-9.

It also follows from the formula shown that to provide the required intensity of X-ray radiation the cathode of the X-ray tube must have a sufficiently great thermionic emission.

The shape and size of the surface of the glass bulb of the X-ray tube are selected with the intention of diffusing all particles that strike its inner surface at a comparatively low temperature. The bulb's glass is chosen in accordance with the wavelengths of the tube's X-ray radiation. In those cases when the X-ray radiation used is very soft and easily absorbed by the glasses of a bulb made of ordinary glass, special kinds of glass are used which have little absorbing effect on soft X-ray radiation, and special areas are made in the bulbs of these glasses, through which the soft X-rays pass.

In essence, an X-ray tube is a two-electrode electronic tube. Therefore, its volt-ampere characteristics are similar to the volt-ampere characteristics of the ordinary diode. Characteristics of spatial charge conditions are approximately subject to the three-halves law, and anode current in saturation conditions is practically determined

only by cathode temperature.

Saturation conditions in X-ray tubes occur at considerably greater anode voltages , by comparison, than in ordinary electronic devices, which is explained by their construction features.

The saturation condition in X-ray tubes is the working condition at which, with constant saturation anode voltage and constant temperature of the cathode, fixed hardness and intensity of X-ray radiation are obtained. Without changing anode voltage, but changing cathode temperature, it is possible to increase or decrease X-ray radiation intensity without changing its hardness, which is very important, for example, in medicine.

To increase intensity and hardness of X-ray radiation, special X-ray tubes are made, which work by pulse operation in which, parallel to cathode overheating, a voltage two or three times greater than nominal is supplied to the tube's anode for several microseconds. Under these working conditions, the tube delivers powerful impulses of X-ray radiation of considerable hardness.

At present, X-ray tubes exist which deliver hard radiation of great power and are intended for inspection of metallic parts and materials of great thickness. Anode voltages of such tubes reach 0.5 to 2 million volts.

16-5. CONSTRUCTION DESIGN FEATURES OF X-RAY TUBES.

X-ray tubes are divided into two classes, according to the method of obtaining free electrons and, partly, by their construction. Ionic X-ray tubes belong to the first class, in which there are always small amounts of gas at pressures of approximately 10^{-3} millimeters of mercury and whose cathodes are not specially preheated. Ionic X-ray tubes are described in the course on ionic devices. At the present time, the

areas of usage of these tubes is limited since they are not entirely perfected.

The electronic X-ray tubes, which we are now considering, belong to the second class; their special features are an incandescent cathode and high vacuum.

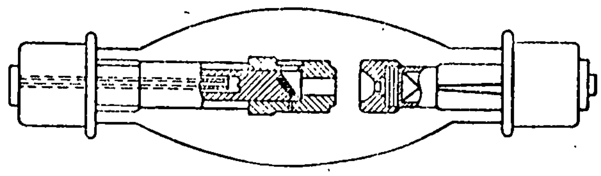

Fig. 16-10. X-ray tube for examining material with flowing oil cooling for work in oil.

Electronic X-ray tubes are divided, according to their usage, into medical and technological tubes.

Medical X-ray tubes, in turn, are subdivided into diagnostic and therapeutic tubes.

Technological X-ray tubes are subdivided into tubes for X-ray inspection of various materials and parts, and tubes for structural and spectrographic analysis.

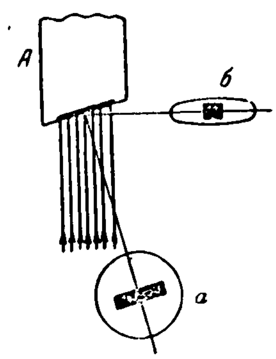

Fig. 16-11. Actual (a) and optical (b) focal points in a tube with line focusing.
A) X-ray tube's anode.

X-ray tubes for structural analysis. Soft X rays, which are strongly absorbed even by ordinary glass, are often used in structural analysis. Therefore, special windows made of "getan"* or vacuum-tight beryllium which have almost no absorption effect on soft X-ray radiation are placed in the glass bulb of tubes of this type for output of X rays. As a result, intensity of soft X-ray radiation is considerably increased. Besides this, radiation intensity is also increased without decrease of focus area, the diameter of which ordinarily is several millimeters.

Since tubes for structural analysis are intended for prolonged work, hollow anodes and water cooling from a pipe line (with a grounded anode) are used in them.

The construction of an X-ray tube for structural analysis is shown in Fig. 16-8. The cathode (1) and the hollow anode (2) are located inside the glass bulb. To protect the bulb from the reflected electrons striking it, the working part of the anode has a sheath (3), in which four holes (4) are made for output of X rays, opposite which four windows (5) are provided in the bulb; for protection from the unused X rays, the anode is connected with a metal cylinder (6), in which four apertures also are provided for opposite the windows. The X rays radiating from the windows permit simultaneous irradiation of four chambers, and the objects placed in them. In tubes of this type, the anodes are cut straight. Targets of various materials to obtain the characteristic rays which are needed for structural analysis are soldered into the surface of this cut.

Special take-down X-ray tubes are made for X-ray spectroemission analysis; the substances under study are placed at their anodes.

X-ray tubes for X-ray inspection of material. The construction of a tube for X-ray inspection of material is shown in Fig. 16-10. The

* [Sic. A type of glass not opaque to soft X rays.]

tube is intended for operation in a protective housing with oil-flow cooling. Some types of X-ray tubes for X-ray inspection of materials attain a length of one meter or more. In such tubes, in addition to electrical focusing which is inadequate because of the great distance between the cathode and the focal point, an additional magnetic focusing by means of a focusing coil, fastened to the tube on the outside, is used. The required focus level of the electron beam is attained by regulating the current flowing through the coil.

Flow cooling of the hollow anode is used in present-day tubes for X-ray inspection of material (especially for screen inspection) to obtain the greatest intensity of X-ray radiation in prolonged work with high anode voltages (up to 150 kv).

To obtain sharp and large images, the focus diameter in some kinds of tubes is carried to several tenths of a millimeter with simultaneous reduction of the maximum permissible power of the tube.

Diagnostic X-ray tubes. Diagnostic X-ray tubes are made for portable apparatuses with working voltages of 60-70 kv and for mobile and stationary apparatuses with a voltage of 75-110 kv. These tubes work for short-term operations on X-ray photographs with a short exposure, but with increased power. However, they do not require intensive cooling, since the heat is slowly dispersed from the focal point to the anode and then to the cooling system during the time of exposure. In a given case, the limits of power permissible for short-term operation of the tube depend only on the size of the focal point and length of exposure, and not on the method of cooling. The greater the time of exposure and the smaller the focal point, the smaller the maximum permissible tube power to obtain the required intensity. Therefore, the majority of present-day diagnostic tubes have radiation cooling.

Increase of power and, consequently, of intensity of X-ray radia-

tion is attained by a line focus, the area of which is approximately three times as great as the area of the visible (optical) focus (Fig. 16-11).

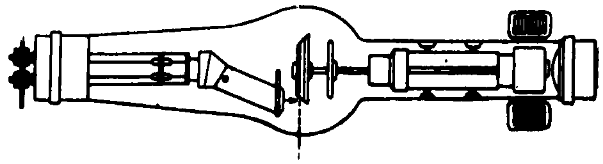

Fig. 16-12. Diagnostic X-ray tube with revolving anode.

In some diagnostic X-ray tubes, there are two focal points and accordingly two cathodes. One of these focal points (the smaller size) is used for X-ray examination and the other (of larger size) for photographs.

The most complicated in construction are the so-called diagnostic tubes with revolving anodes, in which the cathode is offset from the axis of the tube, and a revolving anode having the form of a disc (Fig. 16-12) is fastened at one end of the axis; on its other end a copper cylinder, which is also the rotor of a squirrel-cage motor, is

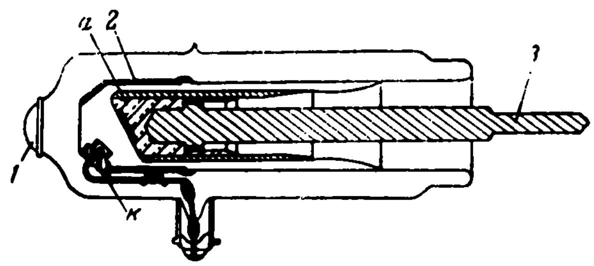

Fig. 16-13. 15-kv X-ray tube for skin therapy.

fastened. The stator of the motor is fastened to the X-ray tube from the outside. When the anode revolves (speeds up to 2700 rpm), the focus is shifted along its surface which remains fixed in relationship to the fixed bulb of the tube. With a constant focal size, such tubes permit, for short work periods, an increase in load by several times normal.

61-102

A line focus is used in tubes with revolving anodes. Thanks to this, with exposures of 0.05 to 0.1 sec and natural anode cooling, tube power reaches 50 kv.

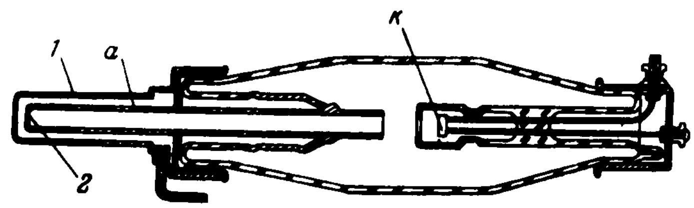

Fig. 16-14. Structure of a therapeutic X-ray tube with hollow anode for body-cavity and short-focus therapy.

Therapeutic X-ray tubes. An example of the construction of a therapeutic X-ray tube for skin therapy at a voltage of up to 15 kv is shown in Fig. 16-13. A window (1) of thin "getan" glass is made in the dome of the bulb for outgoing soft X rays. The tube anode (a) is placed inside a nickel cylinder (2), to which the cathode (K) is fastened. A radiator, which is cooled by a fan, is fastened to the stem of the anode (3), which passes outside the bulb. In a tube of this type the maximum permissible power reaches 150 watts. In operation the cathode of the tube is grounded and the tube is placed in a protective sheath.

The construction of a therapeutic X-ray tube for body-cavity and short-focus therapy is shown in Fig. 16-14. The tube's anode (a) is made in the form of a long copper jacket with a thin bottom and thin sides. The bottom of the jacket extends outside the bulb. A thin-walled brass jacket of larger diameter (1) is pressed onto the outside part of the anode's jacket. A cooling fluid circulates in the space between the walls of the jackets. Electrons radiating from the cathode (K) are focused on the inner surface of the bottom of the copper anode jacket, forming the focal point (2) on it. X rays excited at the focus

61-102

move in all directions, pass through the thin wall of the copper jacket, through the layer of cooling fluid and the thin wall of the brass jacket, and pass outside. The thickness of the walls of the jackets usually amounts to 0.2-0.3 mm. If the inner surface of the bottom of the copper jacket (the anode) is covered with a heavy metal, for example, platinum or gold, then by this means it is possible to increase the intensity of X-ray radiation accordingly.

To insulate the tube from high voltage during operation, it is placed in a protective jacket, and its anode is grounded. Thus, part of the hollow anode is separated from the jacket. The advantage of this tube construction is the possibility of placing it in direct contact with the object being irradiated or inserting it into a cavity of the human body. If it is necessary to separate a narrow beam from the total X-ray radiation in the desired direction a special sheath with an aperture of the required shape is placed on the brass jacket of the tube.

16-6. CONSTRUCTION DESIGN OF SAFETY AND SHIELDED X-RAY TUBES.

In the operation of X-ray tubes, only the part of the X rays included in a small solid angle are used, and the remaining X rays are scattered. These unused rays, like the rays reflected from the objects being studied, are dangerous to operating personnel. Therefore, in work with X-ray tubes, measures are taken for shielding from the unused X-ray radiation. For this purpose, so-called lead chambers, whose walls absorb the unused X rays, are used. In work with tube voltages within the limits of 100 through 400 kv the thickness of the lead wall is, correspondingly, from 1.5 to 15 mm. Along with the lead chambers, which are heavy and bulky and therefore not always convenient, at the present time self-shielded X-ray tubes in which shielding from the

unused X rays is accomplished by means of sheaths for the electrodes inside the tube are used—partial shielding (for example, Fig. 16-10), or by placing the entire bulb of the X-ray tube in special casings—and full shielding (Fig. 16-15). There are special windows in the sheaths and casings for the outgoing X rays being used.

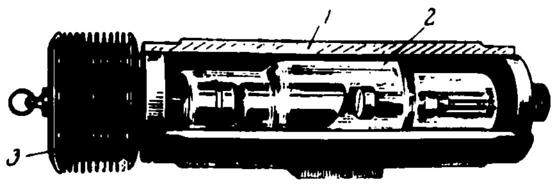

Fig. 16-15. Diagnostic self-shielded X-ray tube with sheath (100 kv).
1) Insulating material with raised surface; 2) X-ray tube; 3) radiator.

The majority of electronic X-ray tubes work in shielded housings, filled with oil to insulate the tube (Fig. 16-16). The dimensions of such tubes are usually smaller than the dimensions of tubes operating in air.

In other types of self-shielded X-ray tubes shielding is accomplished by plating the middle part of the outer surface of the bulb

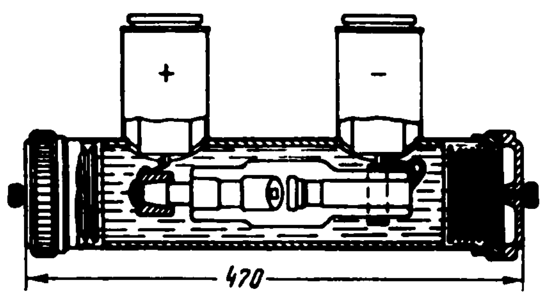

Fig. 16-16. Diagnostic tube, type BDM-100, in shielded casing with oil insulation.

with a layer of lead or by means of making the center part of the bulb of metal (for example, ferrochromium), outside of which a layer

of lead with a window for outgoing rays is placed. Safety X-ray installations, in which the X-ray tube is placed in a metal casing (with mandatory grounding), eliminate the dangers in working with high-voltage supply of X-ray tubes. Besides this, tubes of small dimensions are made, which are, together with a high voltage transformer, placed in oil to increase the electrical resistance of the tube and transformer. In this case, shielding from high voltage and shielding from unused X-ray radiation are both accomplished more easily.

16-7. OPERATION OF X-RAY TUBES. THE X-RAY KENOTRON.

The basic elements of an X-ray installation are the tube and power supply unit — the source of its supply of electrical energy. It is a singlephase high-voltage transformer whose current is rectified by a high-voltage kenotron.

X-ray units are classified according to shielding from high voltage, purpose or usage, and means of transportation.

The first group of installations is open or shielded; the second group, diagnostic, therapeutic, units for structural and spectral analysis, and units for X-ray inspection of materials; the third group, stationary, mobile, and portable units.

An X-ray unit is supplied with test equipment to control tube conditions and a control panel to regulate the incandescence of the tube and kenotron cathodes, to switch on and switch off high voltage, and regulate intensity and hardness of X rays.

Data concerning some types of X-ray apparatus are shown in Table 16-1.

X-ray tubes may work either on alternating or on constant voltage. When the tube is supplied with alternating voltage, it rectifies itself, transmitting current only for a half-cycle.

When the tube is supplied with direct voltage, high voltage of the secondary winding of the transformer is pre-rectified by an X-ray kenotron. Designs for one-kenotron and two-kenotron X-ray units.

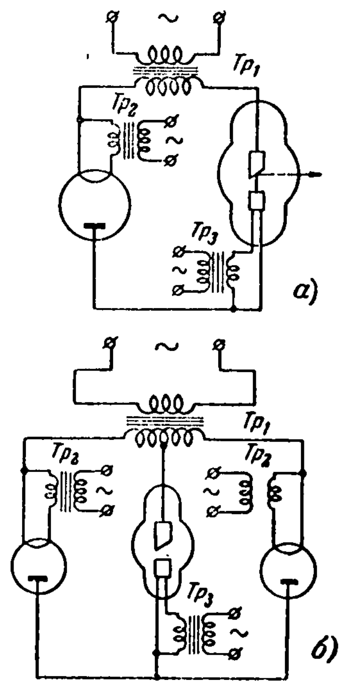

Fig. 16-17. Simplified switch circuit diagram of 1-kenotron (a) and 2-kenotron (b) X-ray units.
1) High voltage transformer; 2) kenotron heater transformer; 3) X-ray tube cathode heater transformer.

exist (Fig. 16-17). The mechanism of conversion of high alternating voltage to direct voltage is usually this: during the positive half-cycle, when the kenotron anode has a great positive potential in relationship to the cathode, electrons move from the cathode to the anode and the kenotron transmits power during this half-cycle. The X-ray tube, in which electrons also move from the cathode to the anode during this half-cycle, also transmits power. In reversed polarity, that is, during the negative half-cycle, when the kenotron anode (and the X-ray tube

TABLE 16-1.
Some types of X-ray equipment.

Типы аппаратов	Назначение	Способ транспортировки	Номинальные значения тока и напряжения: ма	Номинальные значения тока и напряжения: кв	Типы рентгеновских трубок	Типы кенотронов	Количество кенотронов	Наибольшая потребляемая мощность, квт
Д-110-К-4	Диагностика, поверхностная терапия	Стационарный	50	100	ДВ, РДВ	КР-110	4	20
РУ-555	Просвечивание материалов	Передвижной	7	200	БПМ-200	КРМ-110	2	3
РУ-610	Структурный анализ	Переносный	15	50	БСВ-50	—	—	1,5
РУ-560	Диагностика	.	10	65	БДМ-75	—	—	1,2
РУ-720	Диагностика (снимки зубов)	Передвижной	10	65	БДМ-75	—	—	1,2
РУМ-1	Хирургия	.	10	75	БДМ-75	—	—	1,5
РУМ-2	Диагностика, поверхностная терапия	Стационарный	3	100	БДМ-100	—	—	7
УРД-105-К-4	Диагностика, поверхностная терапия	.	3	100	БДМ-100	КГ-110	4	10
РУМ-3	Терапия	.	20	200	ЗБПМ-200	КРМ-150	2	7
РУП-1	Просвечивание материалов	.	20	200	ЗБПМ-200	КРМ-150	2	7

1) Types of equipment; 2) purpose; 3) method of transportation; 4) nominal value of current and voltage; 5) ma; 6) kv; 7) types of X-ray tubes; 8) types of kenotrons; 9) number of kenotrons; 10) greatest required power, kwt; 11) D-110-K-4; 12) RU-555; 13) RU-610; 14) RU-560; 15) RU-720; 16) RUM-1; 17) RUM-2; 18) URD-105-K-4; 19) RUM-3; 20) RUP-1; 21) diagnosis, surface therapy; 22) X-ray inspection of material; 23) structural analysis; 24) diagnosis; 25) diagnosis (X-ray photos of the teeth); 26) surgery; 27) diagnosis, surface therapy; 28) diagnosis, surface therapy; 29) therapy; 30) X-ray inspection of materials; 31) stationary; 32) mobile; 33) portable; 34) ditto; 35) mobile; 36) ditto; 37) stationary; 38) ditto; 39) ditto; 40) ditto; 41) DV, RDV; 42) BPM-200; 43) BSV-50; 44) BDM-75; 45) BDM-75; 46) BDM-75; 47) BDM-100; 48) BDM-100; 49) ZBPM-200; 50) ZBPM-200; 51) KR-110; 52) KRM-110; 53) KG-110; 54) KRM-150; 55) KRM-150.

anode) have great negative potential in relationship to the cathode, the kenotron (and the X-ray tube) do not transmit power, since the kenotron's cold anode (and that of the X-ray tube) does not emit electrons. The so-called inverse or cutoff voltage is established between the cathode and anode of the kenotron, under which condition the kenotron does not conduct current; in modern X-ray kenotrons the amount of this voltage is as much as 400 kv.

In the positive half-cycle, when current flows in the circuit,

a voltage drop is established in the kenotron, the quantity of which is proportional to the internal resistance of the kenotron. If this resistance is great, the voltage drop in the kenotron is also high, and consequently the power developing at the kenotron anode is also great, which causes a powerful heating of the anode and reduces recti-

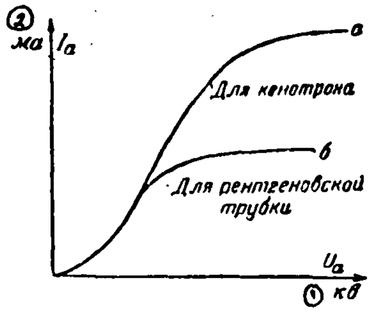

Fig. 16-18. Volt-ampere characteristics of a high-voltage X-ray kenotron and X-ray tube.
a) For the kenotron; b) for X-ray tube.
1) kv; 2) ma.

fier efficiency. Therefore, we strive to make the internal resistance of the kenotron as small as possible, thus reducing the voltage drop in the kenotron and increasing the voltage drop in the load, that is, in the X-ray tube.

The decrease in internal resistance, and consequent drop of voltage in the kenotron, is attained by means of an increase in the working temperature of the kenotron cathode, by which means an emission greater than the emission from the X-ray tube cathode is assured. In other words, the saturation current and the transconductance of the static anode characteristics of the kenotron must be greater than the saturation current and transconductance of the static anode characteristics of the X-ray tube (Fig. 16-18). In order to fill this need, the kenotron cathode is always made more powerful than the X-ray tube cathode. In underheating of the cathode, saturation current and trans-

conductance of the static anode characteristics of the kenotron may become equal to, or even smaller than, the saturation current and transconductance of the static anode characteristics of the X-ray tube. As a result, the internal resistance and voltage drop in the kenotron grow and may become equal to or larger than the internal resistance and voltage drop in the X-ray tube. Therefore, tungsten cathodes, which permit a considerable increase in working temperature and consequently in emission, and tungsten anodes which are capable of dissipating a considerable power output without artificial cooling, are used in X-ray kenotrons. Voltage drop in the kenotron is increased also at the expense of formation of charges on the internal surface of the glass bulb. If the anode is given the shape of a hollow cylinder and the cathode placed inside the cylinder, then the latter prevents electrons from striking the internal surface of the bulb and, consequently, the formation of charges on it.

Voltage drop in an X-ray kenotron under normal cathode incandescence conditions does not exceed 2 kv.

Depending on its purpose, an X-ray kenotron anode is made of nickel, molybdenum, or tantalum. If the cathode temperature of the kenotron is high enough, voltage drop in the kenotron itself is, altogether, only 1-2 kv. X rays are not excited at the kenotron anode under such anode voltages. If the cathode temperature is not high enough, the internal resistance of the kenotron and, consequently, voltage drop in it, will become greater. A great voltage drop in the kenotron causes a considerable acceleration of electrons, which, striking against the anode with great speed, may cause unnecessary and even harmful X-ray radiation. Consequently, kenotron cathode temperature must be sufficient to obtain a great emission, in the presence of which internal resistance of the kenotron is not great.

TABLE 16-2.
X-ray kenotrons.

Обозначения	Условия работы	Предельно допустимое обратное напряжение, кв	Предельно допустимый максимальный анодный ток, ма	Максимальное падение напряжения на кенотроне, кв	Напряжение накала, в	Ток накала, а	Предельно допустимый анодный ток в полупериодной схеме (средние значения), ма		
							0,1 сек	10 сек	Продолжительно
КР-110	В воздухе	110	400	3,0	8—14	7,5—8,7	150	75	30
КР-220	„ „	220	400	3,5	8—14	7,5—8,7	110	55	30
КРМ-110	В масле	110	300	2,5	8—14	7,5—8,7	150	75	30
КРМ-150	„ „	150	150	3,0	8—14	7,5—8,7	110	35	30

1) Designation; 2) working conditions; 3) limit of permissible inverse voltage, kv; 4) limit of permissible maximum anode current, ma; 5) maximum voltage drop in kenotrons, kv; 6) incandescence voltage, v; 7) incandescent current, a; 8) limit of permissible anode current in half-cycle circuit (mean values), ma; 9) 0.1 sec; 10) 10 sec; 11) prolonged period; 12) KR-110; 13) KR-220; 14) KRM-110; 15) KRM-150; 16) in air; 17) in oil.

With a given anode voltage, kenotron cathode temperature must be high enough to obtain a saturation current whose quantity exceeds the current flowing in the X-ray tube. Therefore, in X-ray kenotrons it is suitable to make the cathode of thoriated or carburated tungsten.

Table 16-2 shows types of X-ray kenotrons.

In X-ray kenotrons, as in X-ray tubes, pressure of residual gases must not exceed 10^{-6} through 10^{-7} mm of the mercury column.

There are X-ray instruments which work by pulsed operation. In such apparatuses, a high voltage condenser charge is originated, which is discharged through the X-ray tube, giving a powerful impulse of radiation at this moment.

Since high voltages, reaching tens, hundreds, and thousands of kilovolts, and dangerous for operating personnel, are required for supplying X-ray tubes, control of an X-ray apparatus is accomplished

by a primary (low-voltage) circuit, employing an automatic interlock system for this purpose.

Chapter Seventeen

SEMICONDUCTOR DEVICES

17-1. PROPERTIES OF SEMICONDUCTORS.

A semiconductor rectifier was used in radio receivers in almost the first phase of development of radio engineering. It was a natural semiconductor crystal a (galenite, cassiterite, chalcopyrite and others), which, in contact with a metal catwhisker c (Fig. 17-1) possessed unilateral conduction and served as a detector. A.S. Popov used such a semiconductor detector as early as 1900 to receive telegraph signals from a spark transmitter. Semiconductor detectors occupied a prominent place in the technology of radio reception up until the perfection of vacuum triodes and diodes, when the crystal detector was superseded by the tube.

In 1922 O.V. Losev, a co-worker of the Nizhegorodskaya radio laboratory, in his experiments with semiconductors, having established that a series of crystal detectors possessed falling volt-ampere characteristics, used this effect to generate continuous oscillations and for regenerative reception. The receiver ("Kristadin"),* created by O.V. Losev, was simple, cheap, and required infinitesimal power at small input voltages. O.V. Losev's invention became widely known in the scientific world, but the continuous development and perfection of electronic tubes, as it turned out, entirely superseded the semiconductor detector in practical radio engineering.

In 1948 a semiconductor triode, by means of which it was possible to amplify weak signals, was developed in the USA. Further experiments and theoretical studies permitted not only the discovery of new properties of semiconductors but also the development of new types of semi-

---

* Kristadin - Evidently a brand name.

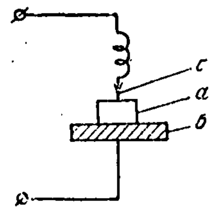

Fig. 17-1. Simplified circuit diagram of a crystal detector.

conductor diodes and triodes.

At the present time, silicon and germanium are chiefly used in making semiconductor diodes and triodes. The atoms of silicon and germanium have four valence electrons in their outer orbits and occupy such a position in a crystal lattice that each of them forms bonds with four adjacent atoms.

Bonds between two atoms are formed by two valence electrons, one of which belongs to one atom and the other electron to another atom (Fig. 17-2). Such a bond is called electron pair or covalent (Fig. 17-3). If the crystal is not subjected to external influences (heat, radiation, or the like), then the bonds in the crystal lattice are not broken and, consequently, there are no free electrons in the crystal, since all valence electrons are in a bound state. In these conditions, the crystal must have the properties of a dielectric. However, for example, at room temperature, the crystal, either of silicon or of germanium, receives heat energy which is enough to displace some electrons from the atoms. Such electrons may be shifted in the crystal. A positive charge, equal in quantity to the charge of an electron, is observed in the place from which an electron has been displaced under the action of external energy. This place may be occupied by an adjacent electron, and the place of the adjacent electron by the next electron, and so forth. In other words, the place which lacks an electron may be independently shifted in the crystal, which is equivalent to displacement of a

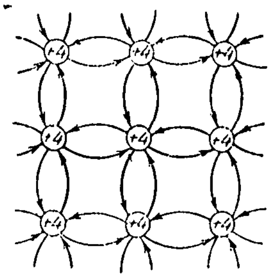

Fig. 17-2. Surface diagram of a crystal of silicon (or germanium). The number of positive charges, corresponding to the number of valence electrons, is shown in the circles.

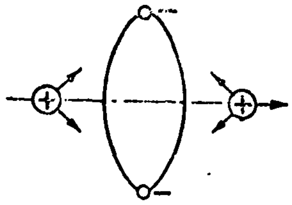

Fig. 17-3. Model of the hydrogen molecule. Electrons moving along the common orbit compensate for the force of repulsion betwen the nuclei.

particle having a positive charge equal in quantity to the charge of an electron.

Thus, breaking a monovalent bond in a semiconductor crystal produces two types of charged particles: an electron and a place where an electron is lacking, which has been called a "hole." If there is no external electrical field, both electrons and holes are shifted at random (chaotically) in the crystal as a consequence of thermal motion. In the presence of an external field, motion of electron and hole becomes orderly and electrical current originates in the crystal.

Conductions created by directional shifting of electrons and holes have been named electron and hole conduction and are designated accordingly as N-type conduction (N meaning negative) and P-type conduction (P meaning positive). In metals, holes are neutralized practically instantaneously by electrons and consequently do not participate in conduction. In semiconductors holes are not neutralized so quickly as in metals, as a result of which they may be shifted and take part in conduction. It follows that hole conduction is different from ion conduction; in ion conduction the ions themselves are shifted directly, but in hole conduction electrons are shifted, alternately

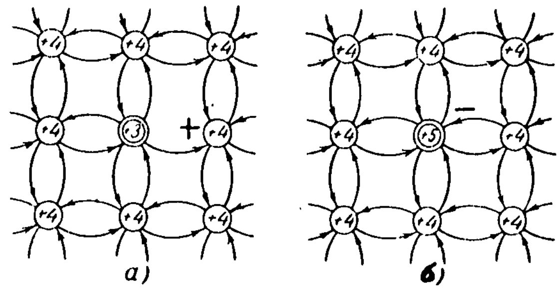

Fig. 17-4. Surface diagram of a crystal of silicon (or germanium). a) With an acceptor impurity, forming a hole-type (P-type) conduction; b) with a donor impurity, forming an N-type conduction.

shifting from bond to bond and neutralizing positive charges (holes) where they move but forming them in the places from which they move. Therefore, the nuclei of the atoms remain stationary, but positive charge conditions (holes) are shifted in directions opposite to the directions of electron movement. In a semiconductor crystal of maximum purity (without impurities), the number of free electrons is equal to the number of holes, since each electron freed by external influences forms one hole. Therefore, equal amounts of free electrons and of holes formed as a result of the electrons leaving their bonds take part in electrical conduction. Conduction accomplished simultaneously by directional shifts of equal numbers of electrons and holes has been named the intrinsic conductance of a semiconductor. Thus, a semiconductor which is free from impurities may have electron and hole conduction simultaneously. It is clear that such semiconductors cannot rectify alternating current, since in the presence of an external alternating field the current in these semiconductors will flow with any field alignment.

We will now consider a semiconductor having small impurities of other substances in its volume. The atoms of the impurities may be di-

vided into two types according to the number of valence electrons; atoms which have fewer valence electrons than the atoms of the semiconductor, for example, three-valence atoms, and atoms which have more valence electrons than the atom of the semiconductor, for example, five-valence atoms. A three-valence atom of the impurity takes the place of a semiconductor atom in the crystal lattice and is surrounded by four atoms of the semiconductor.

By this means, three valence electrons of the atom of the impurity are coupled with the electrons of three surrounding atoms of the semiconductor. A hole is formed in the bond with the fourth of the adjacent atoms, since the electron of this atom has no pair among the electrons of the atom of the impurity. In the presence of sufficient external energy, this hole may be shifted along the crystal and participate in conduction. The atom of the impurity does not participate in conduction, being turned into a negative ion after leaving the hole (Fig. 17-4a).

Impure semiconductors in which hole conduction takes place are called hole semiconductors (P-type), and impurities whose atoms accept electrons from semiconductor atoms in their bonds are called acceptors.

If a five-valence atom of the impurity takes the place of a semiconductor atom, it forms four bonds with adjacent semiconductor atoms, and the fifth valence electron, which under these conditions appears to be weakly bound with the atom of the impurity, breaks loose from its atom by thermal motion energy. The energy required to break this fifth electron loose is insignificant; for antimony and bismuth, for example, it amounts to about 0.01 ev. The electrons lost by the atoms of the impurity become free and participate in conduction (electron conduction and the atoms of the impurity itself become fixed positive ions, each of which, as in the atoms of the semiconductor, has four valence electrons (Fig. 17-4b).

Impure semiconductors in which electron conduction takes place are called electron semiconductors (N-type), and impurities whose atoms release their valence electrons and thus provide electron conduction are called donors. Five-valence phosphors, arsenic, and antimony are often used as donor impurities, and three-valence aluminum, boron, indium and galium as acceptors.

Thus, if there are donor impurities in the crystal of silicon or germanium, it has electron conduction. In the presence of acceptor impurities in the crystal hole conduction takes place.

Electrons in N-type semiconductors and holes in P-type semiconductors are called majority carriers of current, and the small amount of electrons in P-type semiconductors and of holes in N-type semiconductors are minority carriers of current.

Minority carriers have a limited lifetime. However, during this time they are subjected to numerous collisions with the atoms of the lattice and in the presence of an external field participate in conduction. Minority carriers play an important role in the mechanism of semiconductor devices operation.

As we have indicated, elements of group IV, germanium and silicon, are used to make semiconductor devices. However, many compounds of the elements of groups III and V (the so-called intermetallic compounds) have properties of semiconductors resembling the properties of the elements of group IV. Among these compounds are included, for example, the antimonides of indium (InSb), gallium (GaSb) and aluminum (AlSb), the phosphides of indium (InP), gallium (GaP) and aluminum (AlP) and the arsenides of indium (InAs), gallium (GaAs) and aluminum (AlAs). It is suspected that some of the intermetallic compounds may turn out to be more suitable materials (both in the technology of their processing and in their properties) for production of semiconductor devices than ger-

manium and silicon. However, they have not yet been sufficiently studied.

Electron and hole impurity semiconductors having impurity conduction, that is, primarily either only electron or only hole conduction, like semiconductors of maximum purity having electron and hole conductions simultaneously and at an equal level, cannot rectify alternating current, since current will flow through such semiconductors at any polarity of input voltage.

17-2. PRINCIPLES OF OPERATION OF SEMICONDUCTOR DIODES AND TRIODES.

We will consider flow phenomena on the boundary of contact of two semiconductors having different conductions, using as an example a semiconductor which consists of three regions: the N-region in which the majority carriers of electrical charges are electrons created by a donor impurity, the P-region in which the majority carriers are holes created by an acceptor impurity, and the region separating them which is called the electron hole junction region (N-P junction), in which acceptor and donor impurities are maintained approximately equally. Free electrons located in the electron region of the semiconductor, N, may diffuse into the hole region, P, in which there are two free electrons. As a result of such diffusion, the hole region of the semiconductor is negatively charged, and the electron region positively (Fig. 17-5a).

On the other hand, holes may diffuse from the hole region, where there are many of them, into the electron region, where there are few. Diffusion of holes also leads to the hole region being additionally negatively charged and the electron region positively. As a result, negative charges are accumulated on one side of the junction region and positive charges on the other. Thus, as a consequence of diffusion

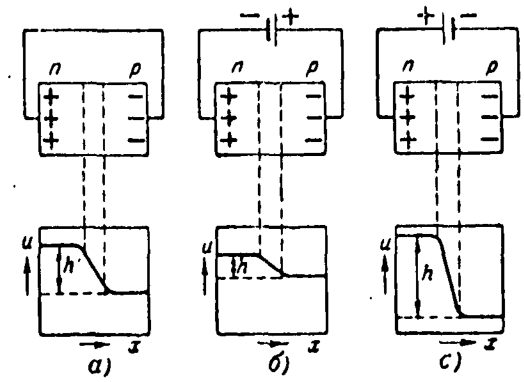

Fig. 17-5. Change of height of potential threshold in a semiconductor with P-N junction under the action of external potential differences.
a) State of equilibrium;
b, c) decrease and increase of potential threshold by means of external potential differences.

of majority carriers of charges (electrons and holes) between the electron and hole regions in reciprocally opposite directions, some internal difference of potential (contact potential difference) is established. This prevents further diffusion of electrons and holes, that is, it represents a potential threshold for them. Therefore, only a small part of the majority carriers of charges (electrons and holes), which have great enough energies, may overcome the retarding field of potential threshold and transit from their region into that opposite. The higher the potential threshold, the fewer the number of major carriers of charges that will be able to overcome it and transit from their region into the other one. However, a minor number of electrons in the P-region and holes in the N-region (minority carriers of charges) do not encounter any impediments in movement across the potential threshold, since the field of the potential threshold is an accelerating field for them. Therefore, the movement of electrons from the P-region into the N-region and of holes from the N-region into the P-region is

not dependent on the height of the potential threshold. Thus, electrons and holes in a semiconductor with electron-hole junction move both from the N-region to the P-region and in the opposite direction. Hence, when there is no external voltage, current is not observed in the circuit, since the flow of electrons from the N-region and of holes from the P-region is equal to the reverse flow of electrons from the P-region and of holes from the N-region, that is, a state of equilibrium is established, which is characterized by a fixed height of potential threshold $\underline{h}$, in the presence of which an equilibrium of the above-mentioned major and minor carriers of electrical charges takes place.

If we now apply a minus external voltage to the N-region and a plus external voltage to the P-region from the power supply (Fig. 17-5b), the applied difference of potential becomes opposite in sign to the contact difference of potential. As a result, the height of potential threshold $\underline{h}$ is lowered and the state of equilibrium is disrupted, since the lower the height of the potential threshold, the more electrons will transit from the N-region to the P-region and the more holes from the P-region to the N-region, but the number of electrons and holes moving in the opposite direction remains the same, since this is not dependent upon the height of the potential threshold. Thus, current, whose quantity is equal to the sum of the electron current flowing from the N-region to the P-region and of the hole current flowing from the P-region to the N-region, appears in the external circuit. With an increase of the external voltage, the height of the potential threshold is decreased, and the current in the external circuit is increased.

If we change the polarity, applying positive voltage to the N-region and negative to the P-region (Fig. 17-5c), the applied difference of potentials will become the same sign as the contact difference of potentials and the height of the potential threshold will increase.

As a result of this, the flow of major charge carriers from one region to the other is decreased, and the reverse flow of the minor charge carriers remains the same, since it is not dependent on the height of the potential threshold. By this means, the quantity of current created by the transitions of the major carriers may become less than the current created by the transitions of the minor carriers, and a reverse current will appear in the external circuit. If the external voltage is increased, the potential threshold is increased, the flow of the major carriers is decreased, and the flow of the minor carriers remains unchanged but the flow of reverse current in the circuit is increased. At some difference of potential, the flow of the major carriers of charges will become practically equal to zero, and the reverse current will attain maximum value. As a result, a small quantity of saturation current will be established in the external circuit.

Thus, with a changed polarity, the external voltage raises the potential threshold and it is as though a valve were shut in the path of the flow of the major carriers of electrical charges. Therefore, the current in the external circuit is very small (the circuit is cut off). Naturally, if alternating voltage is applied to a semiconductor with electronic-hole conduction, the current will flow in the external circuit only in the half-cycles during which the N-region has a negative potential and the P-region a positive potential. Consequently, such a semiconductor may be used as a rectifier of alternating current (a semiconductor diode).

The principles of operation of a semiconductor diode which have been described are applicable to a diode with junction contact in which electron-hole junctions are used. The principle of operation of a point-contact diode, in which rectification is accomplished in a thin layer, nascent in the crystal under a contacting metallic catwhisker, is not

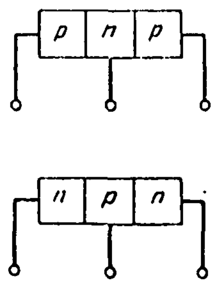

Fig. 17-6. Schematic diagram of P-N-P and N-P-N types of junction crystal triodes.

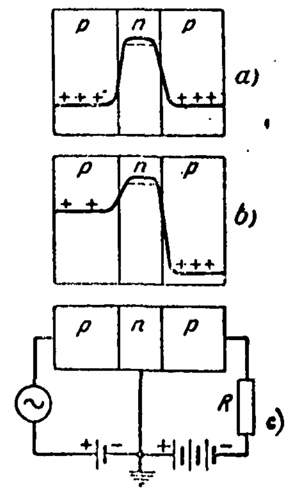

Fig. 17-7. Change of height of potential threshold in a P-N-P triode.

essentially different in principle of operation from a diode with junction contact, since this layer takes the place of the electron-hole junction. The names "point-contact diode" and "junction diode" do not so much emphasize differences in working principles of these diodes as their details of construction.

We will now consider the processes in semiconductors with electron-hole junctions used in junction triodes.

In a junction triode, in place of the one N-P junction present in a diode of this type, there are used two junctions of the N-P-N type, in which the center region has hole conduction and the side regions electronic conduction, or of the P-N-P type, when the center part has electronic and the side regions hole conduction (Fig. 17-6). One of the junctions is switched into the circuit in a forward direction (the emitter circuit), and the other in a cutoff direction (the collector circuit). The center region of the junction is switched in as a base (or foundation) and is made very thin (from several microns to several tens of microns).

Without external voltage the potential distribution in a triode with P-N-P junctions corresponds to the curve in Fig. 17-7a, showing that the height of the potential threshold at the junction is the same. If the external voltage is switched on, the potential threshold at the junction switched on in a forward direction is decreased, and the potential threshold at the junction switched on in a cutoff direction is increased (Fig. 17-7b). By this means, holes in great quantities advance from the emitter to the base and continue their movement in the base primarily by diffusion, since the voltage drop in the thin layer of the base is insignificant. At their exit from the base the holes advance into the acceleration field and fall on the collector.

With the change of current in the emitter circuit, the number of holes striking the collector, and, consequently, the current in the collector circuit also, is changed. But the change of current in the collector circuit is somewhat smaller than the change of current in the emitter circuit. This is explained by the fact that not all the holes which have left the emitter reach the collector. During the transition from the emitter to the collector part of the holes recombine with electrons. At best, when the distance between the emitter and the collector is small, the base thinnest, and, consequently, the recombination speed is small, 99.8 per cent of the holes which have left the emitter reach the collector. Therefore, in junction triodes the amplification factor of the current, which is a relationship of the quantity of current flowing in the collector circuit to the quantity of current flowing in the emitter circuit, is somewhat less than one (from 0.9 to 0.98).

The amplifying properties of a junction triode also are extremely dependent on the unevenness of thickness of the base. If, for example, the base has a concavo-concave shape, the path and, consequently, the

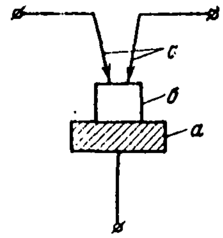

Fig. 17-8. Point-contact triode.
a) Metal base; b) germanium crystal; c) metal catwhisker.

transit time of these paths by holes will be different. As a consequence of this, holes which have left the emitter simultaneously will strike the collector at different times, the signal in the collector circuit is drawn out in time, the amplitude of the current in the collector circuit becomes less than the amplitude of the current in the emitter circuit and the amplifying properties of the triode deteriorate. The higher the frequency of the amplified signal, the greater the distance between the emitter and the collector, and the thicker and more uneven the base, the greater the decrease of current amplitude in the collector circuit. However, the manufacture of too thin a base is fraught with great technological difficulties and is not always expedient, since if the base is too thin the feedback from the collector to the emitter is amplified.

To compute the amplification factor of a junction triode from power and voltage, let us presume, for simplicity, that practically all the holes transiting from the emitter to the base fall on the collector. Therefore, if the current in the emitter circuit is changed by any quantity, the current in the collector circuit is changed by the same quantity. Under these conditions, the power consumed in the emitter circuit by change of current i by quantity Δi and voltage u_e* by

---

* [$u_э$ - u_e - $u_{emitter}$.]

quantity Δu_e* and power yielded in the collector circuit at load R at a change of current by the same quantity Δi, will, accordingly, be equal to: W_e** = $\Delta u_e \Delta i$ and W_n*** = $\Delta i^2 R = \Delta i \Delta u_n$, where Δu_n**** is change of voltage drop at load R for change of current by quantity Δi. Knowing the power consumed in the emitter circuit and the power yielded at the load of the collector circuit, amplification factor may be determined from power and voltage:

$$\frac{W_н}{W_э}=\frac{\Delta i^2 R}{\Delta u_э \Delta i}=\frac{\Delta i R}{\Delta u_э}=\frac{\Delta u_н}{\Delta u_э}.$$

In the given case, the current amplification factor α is equal to one. Increase of α up to ten is attained in special hookups.

Some types of germanium triodes have power and voltage amplification factors as high as several thousand.

If the sign of all currents and voltages is changed, the principles of operation described for a triode with P-N-P junctions will also be correct for a triode with N-P-N junctions. With proper supply polarity, one catwhisker, switched into the circuit of a crystal triode, "emits" a flow of holes into the volume of a crystal (directed toward the base) which has electron conduction. In reverse polarity and hole conduction of the crystal, the catwhisker emits a flow of electrons into the volume of the crystal. Therefore, this whisker is called the emitter. The potential of the other whisker is opposite in sign to the charges emanated from the emitter into the volume of the crystal. If the emitter emanates electrons, the second whisker must have a positive potential; if holes, it must have negative potential. Thus, in accordance with polarity, a flow of electrons or holes emanated by the emitter will

---

* [$u_э$ - u_e - $u_{emitter}$ - $u_{emitter}$.]
** [$W_э$ - W_e - $W_{emitter}$ - $W_{emitter}$.]
*** [W_H - W_n - $W_{nagruzka}$ - W_{load}.]
**** [u_H - u_n - $u_{nagruzka}$ - u_{load}.]

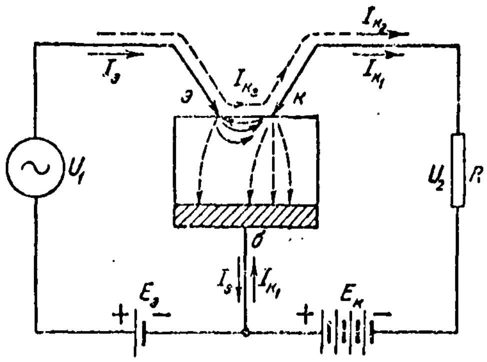

Fig. 17-9. Simplified circuit diagram of a point-contact triode. 1) Emitter; 2) collector; 3) variable signal voltage; 4) amplified signal voltage; 5) base. Trajectories of electrons in the crystal are shown by dotted lines; trajectories of holes, by solid lines.

fall on the second whisker. Consequently, the second whisker fulfills the function of a collector. The quantity of the flow of holes or electrons emanated by the emitter changes with changing potential of the triode base (the metallic plate) in regard to the emitter.

A simplified circuit diagram of a germanium point-contact triode, having electron conduction, consists of the emitter circuit and the collector circuit (Fig. 17-9). A source of constant voltage E_e* is switched into the emitter circuit e, giving the emitter a small positive bias in relation to the triode base, and, in series with this, the alternating signal voltage U_1, subject to amplification. The direct current flowing in the emitter circuit I_e** (of the order of 1 ma), the direction of which is indicated by the arrow, corresponds to the direction of forward current in the contact, that is, to a forward direction. The resistance of load R is switched into the collector circuit c,

* [$E_э$ - E_e - $E_{emitter}$ - $E_{emitter}$.]
** [$I_э$ - I_e - $I_{emitter}$ - $I_{emitter}$.]

from which an amplified signal voltage is taken off; this is the source of the voltage E_k,* whose polarity must be such that the collector has a negative voltage in relationship to the base and the emitter. The direction of the direct current in the collector circuit I_{k1}** is shown by the arrow and corresponds to the direction of the reverse current in the collector contact.

The resistance of the barrier layer, created in the crystal at the indicated polarity in the collector circuit, is considerably greater than the resistance to the forward current in the crystal. Therefore, the current I_{k1}, flowing in the collector circuit at $E_k \approx E_e$, is very small. A current of 1 ma may flow at a supply voltage E_k equal to several tens of volts. Thus, the current amplification factor α does not exceed one.

Actually, the current amplification factor of point-contact triodes is considerably greater than one. This is explained by the fact that, besides current I_{k1}, current I_{k2}*** flows in the same direction; it is created by the drift of a considerable number of holes from the emitter to the collector (the flow of holes from the emitter to the base is insignificant, since the distance between the emitter and the base is many times greater than the distance between the emitter and the collector).

The over-all current $I_{k1} + I_{k2}$ changes in quantity with the frequency of current change in the emitter circuit, and, flowing in the collector circuit, creates an alternating voltage drop at the load resistance. The greater the amount of alternating voltage U_2 taken

---

* [E_k - E_k - $E_{kollektor}$ - $E_{collector}$.]

** [I_{k1} - I_{k1} - $I_{kollektor\ 1}$ - $I_{collector\ 1}$.]

*** [I_{k2} - I_{k2} - $I_{kollektor\ 2}$ - $I_{collector\ 2}$.]

from this resistance in comparison with the alternating voltage of the signal U_1, the greater the triode's amplification factor according to voltage. Since the resistance of the emitter circuit is insignificant, then a minor change of voltage in this circuit noticeably changes the quantity of current. In the collector circuit, however, resistance is great. Therefore, a minor change of current in the collector circuit considerably changes the voltage drop at load resistance R. Thus, point-contact triodes, similar to vacuum triodes, may operate as amplifiers.

17-3. CONSTRUCTION OF SEMICONDUCTOR DIODES AND TRIODES.

A point-contact diode consists of a semiconductor crystal and a metallic pointed catwhisker. Germanium and silicon are used as crystals, and the catwhisker is made of tungsten or phosphorous bronze of a diameter of about 0.1 mm (Fig. 17-10). The point of the metallic catwhisker (3) is pressed against the polished surface of the crystal (1). The crystal is brazed to the metallic crystal holder (2). The catwhisker and crystal are located inside a ceramic tube (5) between metallic flanges (4). Leads (6), by which the diode is soldered into the circuit, are led from the crystal and catwhisker. The crystal end of the catwhisker is sharply pointed. The area of contact of the catwhisker with the crystal usually amounts to several square microns, with a total polished crystal surface of the order of 1 mm^2.

A point-contact triode consists of a crystal, usually germanium, with electron conduction, and two metallic catwhiskers, of which one serves as an emitter and the other as a collector (Fig. 17-11a). The crystal (1) is brazed to the metallic base (2), which serves as a control electrode. Two metallic catwhiskers of tungsten or phosphorous bronze are pressed against the free surface of the crystal. The dis-

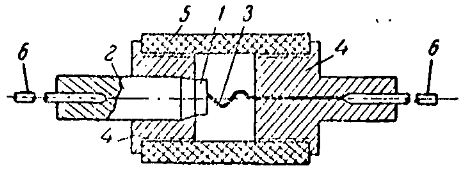

Fig. 17-10. Construction design of a point-contact crystal diode.

tance between the contact points of the catwhiskers amounts to 0.05-0.25 mm. There are two leads (4) from the catwhiskers. The base (2) serves as a third lead. The triode is switched into the circuit by these three leads. The leads from the emitter and collector pass through the insulating plug (5). The base (2), crystal (1), and plug (5) with the lead (4) are mounted in a metallic sleeve (6).

A junction triode is shown in Fig. 17-11b. A crystal of germanium (1) is fastened into a metallic crystal holder (2); the collector (4) and the emitter (3) are located on its two sides. Three-valence indium serves as a material for the collector and emitter electrodes, to which the leads (7) are brazed. The crystal holder is fastened into a metallic case (5). The leads from the emitter and collector (7) are insulated from the case by glass insulation (6). The lead from the crystal (8) is connected with the case.

Together with the semiconductor devices we have described, the so-called semiconductor photodiodes or, as they are called, phototransistors, in which the current flowing across the semiconductor junction

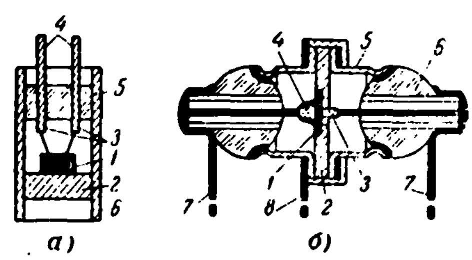

Fig. 17-11. Construction of point-contact (a) and junction (b) triodes.

is changed by the action of light, are also widely used. In construction, a phototransistor is a semiconductor diode; however, in operating principles it is like semiconductor triodes, since the light which strikes it operates in a similar manner to the operation of the emitter in a crystal triode.

The operating principle of the phototransistor is based on the fact that the electrons of a semiconductor increase their energy and break their valence bonds under the action of light. Such electrons may participate in conduction. Holes are also formed simultaneously with the electrons. If some potential difference is applied to the semiconductor junction, the concentration of free current carriers in the semiconductor is increased under the action of light, its resistance is decreased, and the current flowing across the junction increases. Thus, with a change of the flow of light, the current in a phototransistor circuit is also changed.

Phototransistors may be made either as P-N junction transistors or as point-contact transistors, since it is not mandatory to subject the P-N junction to the action of light, but either of the P-regions or N-regions. To decrease the influence of recombination, the drift time of the current carriers is decreased by decreasing the width of the P- and N-regions.

Phototransistors are used in various devices; they are very practical, since they are small in dimensions and weight, and have great sensitivity. The efficiency of phototransistors is increased by decreasing the wavelength of light, since by this means the photon energy is increased. However, at too great a photon energy, that is, at very short wavelengths, the electrons acquire too great an energy and may transit outside the semiconductor. As a result, photoelectron emission may be excited. Therefore, phototransistors are not used for

work in the blue and ultraviolet regions of the spectrum.

17-4. PARAMETERS AND CHARACTERISTICS OF SEMICONDUCTOR DIODES.

Junction diodes transmit considerably greater currents and diffuse greater power at the contact in comparison with point-contact diodes, which is why they are used as powerful rectifiers.

The properties of semiconductor diodes, like those of vacuum diodes, are determined by their parameters and characteristics. The system of parameters used for junction diodes differs from the system of parameters used for point-contact diodes, since measurement of junction diodes is conducted chiefly on alternating current, and of point-contact diodes, on direct current.

The values of the basic parameters for some germanium diodes of point-contact and junction type are shown in Tables 17-1 and 17-2 respectively.

The diodes are characterized by these parameter values in operation in half-wave rectification circuits, with an active resistance load (without capacitance).

Besides the above-mentioned parameter, the interelectrode capacitance of semiconductor diodes is also an important parameter. In point-contact diodes this capacitance is insignificant (from 0.2 to 0.4 $\mu\mu f$), and in junction diodes it reaches 50 $\mu\mu f$. Therefore, point-contact diodes work on higher frequencies.

The characteristics of semiconductor diodes give more complete information about their properties for various purposes.

The static volt-ampere characteristics, expressing the relationship of forward current to input voltage, is the basic. The volt-ampere characteristics of the worst (curve 1) and of the best (curve 2) point-contact germanium diodes of type DG-Ts8 are shown in Fig. 17-12. These

TABLE 17-1.

1) Диоды	2) Наибольший прямой ток, *ма*	3) Наибольший обратный ток, *ма*	4) Наименьший прямой ток при напряжении 1 *в*, *ма*	5) Наибольшее обратное рабочее напряжение, *в*	6) Наименьшее обратное пробивное напряжение, *в*
7) ДГ-Ц5	16	11) 0,25 при—70 *в*	1	75	100
8) ДГ-Ц6	16	12) 0,8 при—100 *в*	2,5	100	125
9) ДГ-Ц8	24	13) 0,5 при—30 *в*	10	30	50
10) ДГ-Ц12	16	14) 0,5 при—10 *в*	5	30	45

1) Diodes; 2) greatest forward current; ma; 3) greatest reverse current, ma; 4) least forward current at a voltage of 1 v, ma; 5) greatest reverse working voltage, v; 6) least reverse breakdown voltage, v; 7) DG-Ts5; 8) DG-Ts6; 9) DG-Ts8; 10) DG-Ts12; 11) 0.25 at 70 v; 12) 0.8 at 100 v; 13) 0.5 at 30 v; 14) 0.5 at 10 v.

TABLE 17-2.

1) Диоды	2) Выпрямленный ток, *ма*	3) Обратный ток, *ма*	4) Подводимое переменное напряжение (эфф.), *в*	5) Прямое падение напряжения, *в*
6) ДГ-Ц21	300	⩽0,5	35	⩽0,5
7) ДГ-Ц22	300	⩽0,5	70	⩽0,5
8) ДГ-Ц23	300	⩽0,5	105	⩽0,5
9) ДГ-Ц24	300	⩽0,5	140	⩽0,5

1) Diodes; 2) rectifier current, ma; 3) reverse current, ma; 4) direct input voltage (effective), v; 5) forward voltage drop, v; 6) DG-Ts21; 7) DG-Ts22; 8) DG-Ts23; 9) DG-Ts24.

characteristics show that semiconductor diodes may serve to rectify alternating current.

Besides the volt-ampere charactistics, frequency characteristics, showing the relationship of the rectified diode current to the frequency of the alternating supply voltage, are most important for a point-contact germanium diode. Typical frequency characteristics of a diode of the DG-Ts type for two values of load resistance R = 1 kohm

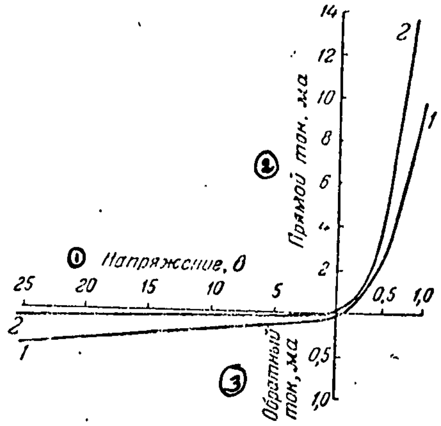

Fig. 17-12. Volt-ampere characteristics of type DG-Ts8.
1) Voltage, v; 2) forward current, ma; 3) reverse current, ma.

and R = 100 kohm are shown in Fig. 17-13. It is not difficult to see that, in the first place, rectified current is decreased with increase of frequency of supply voltage and, in the second place, as far as increase of load resistance goes, rectified current depends on frequency to a lesser degree. In the given case, the frequency limit exceeds 150 Mc, since point-contact germanium diodes have a small transfer capacitance. Junction diodes as a consequence of their greater transfer capacitance are not used on frequencies above 50 kc.

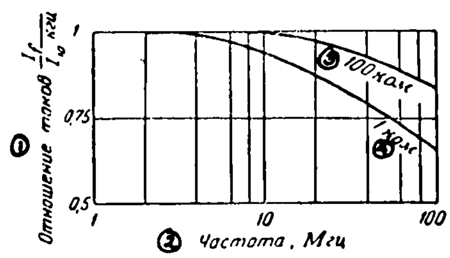

Fig. 17-13. Frequency characteristics of a point contact germanium diode for two values of load resistance.
1) Relationship of currents I_f/I_{10} kc; 2) frequency, Mc; 3) 100 kohm; 4) 1 kohm.

17-5. PARAMETERS AND CHARACTERISTICS OF SEMICONDUCTOR TRIODES.

The resistances characterizing a triode in selected operating

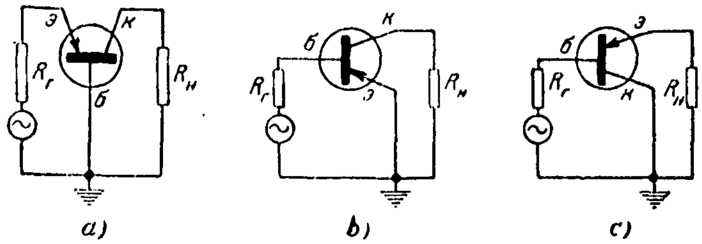

Fig. 17-14. Current diagrams of triodes. a) With grounded base; b) with grounded emitter; c) with grounded collector.

conditions are the basic parameters of a point-contact triode. These resistances include input and output resistance as well as feedback resistance.

Semiconductor triodes operate in three basic hookups: with grounded base, grounded emitter, or grounded collector (Fig. 17-14). Junction triodes operate successfully in all three hookups. Point-contact triodes operate stably only in a hookup with a grounded base. In this case, the resistances enumerated above are determined by the following means:

The input resistance r_{11}, which is the sum of the resistances of the emitter r_e* and of the base r_b** with an open output: $r_{11} = r_e + r_b$.

The output resistance r_{22} (the sum of the resistances of the collector r_k*** and of the base r_b with an open input): $r_{22} = r_k + r_b$.

Feedback resistance r_{12} (resistance of the base r_b at open input): $r_{12} = r_b$.

By means of the resistances r_{11}, r_{22}, and r_{12}, it is possible to determine resistances r_e, r_k, and r_b separately: $r_e = r_{11} - r_{12}$; $r_k = r_{22} - r_{12}$; $r_b = r_{12}$.

Since a circuit diagram with a common base (Fig. 17-14a) is or-

---

* [$r_э$ - r_e - $r_{emitter}$ - $r_{emitter}$.]

** [$r_б$ - r_b - r_{baza} - r_{base}.]

*** [r_k - r_k - $r_{kollektor}$ - $r_{collector}$.]

dinarily used to read the characteristics of a triode, the values of resistances r_{11}, r_{22}, and r_{12} may be determined directly from the characteristics:

$$r_{11} = \left| \frac{\Delta u_э}{\Delta i_э} \right|_{i_к = \mathrm{const}};$$

$$r_{22} = \left| \frac{\Delta u_к}{\Delta i_к} \right|_{i_э = \mathrm{const}};$$

$$r_{12} = \left| \frac{\Delta u_э}{\Delta i_к} \right|_{i_э = \mathrm{const}}.$$

The so-called static current amplification factor α is also an important parameter of a triode:

$$\alpha = \left| \frac{\Delta i_к}{\Delta i_э} \right|_{u_к = \mathrm{const}}.$$

Besides the parameters enumerated above, a point-contact triode is characterized by a noise factor, maximum dissipated power at the collector, cutoff amplification frequency, and power amplification factor, expressed as the relationship of oscillation power P_k* discharged at the load of the triode R_n,** to the nominal signal supply power P_n***:

$$\frac{P_к}{P_н} = \frac{u^2_{вых}}{R_н} : \frac{u^2_r}{4R_r} = \frac{4u^2_{вых}R_r}{u^2_r R_н},$$

where u_{vykh} is alternating component of output voltage; u_r is emf of input signal supply; and R_r is internal resistance of input signal supply.

The basic parameters of junction triodes include the static current amplification factor, power amplification factor, efficiency, and the noise factor. For comparison, some parameters of point-contact and

---

* [P_k - $P_,$ - $P_{kollektor}$ - $P_{collector}$.]
** [R_H - R_n - $R_{nagruzka}$ - R_{load}.]
*** [P_H - P_n - $P_{nagruzka}$ - P_{load}.]

TABLE 17-3.

1) Параметры	2) Вакуумные триоды	3) Полупроводниковые 4) точечные	5) плоскостные
6) Коэффициент полезного действия в режиме класса А, %	1—25	30—35	45—49
7) Коэффициент полезного действия в режиме класса В, %	20—50	18—30	30—50
8) Граничная частота, *Мгц*	>100	30—50	3—5
9) Коэффициент шума, *дб*	5—15	45	15
10) Предельная температура, °C	500	70	70
11) Предельный срок службы, *ч*	5 000	70 000	90 000
12) Минимальная потребляемая мощность, *вт*	0,1	10^{-3}	10^{-6}
13) Коэффициент статического усиления по току	—	2—4	0,9—0,98
14) Мощность рассеяния на коллекторе, *вт*	—	0,02—0,2	18) До 20
15) Ток эмиттера, *а*	—	0,001—0,04	19) До 5—10
16) Ток коллектора, *а*	—	0,002—0,04	20) До 5—10
17) Напряжение на коллекторе, *в*	—	20—100	50—200

1) Parameters; 2) vacuum triodes; 3) semiconductor; 4) point-contact; 5) junction; 6) efficiency in class A operation, per cent; 7) efficiency in class B operation, per cent; 8) cutoff frequency, Mc; 9) noise factor, db; 10) limiting temperature, degrees C; 11) limiting lifetime, hours; 12) minimum power requirement, watts; 13) static current amplification factor; 14) dissipation of power at collector, watts; 15) emitter current, a; 16) collector current, a; 17) voltage at collector, v; 18) to 20; 19) to 5-10; 20) to 5-10.

junction triodes, as well as vacuum triodes, are shown in Table 17-3.

Comparison of the data in this table shows, in the first place, that semiconductor triodes have a number of substantial advantages in comparison to vacuum triodes. For example, they require an extremely small amount of power and have a long, useful life. One should add to these their small dimensions and weight and also great mechanical strength. Besides this, junction triodes, in comparison to point-contact triodes, have a greater mechanical strength thanks to the lack of point-contacts, as well as considerably less intrinsic noise and greater output power.

However, semiconductor triodes also have a number of substantial shortcomings, among which should, first of all, be noted their small

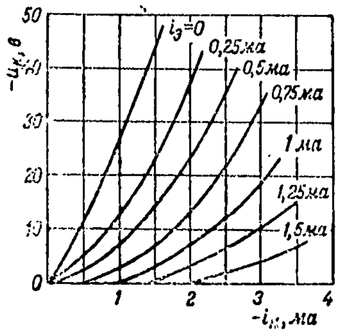

Fig. 17-15. Static output characteristics of a point-contact triode.

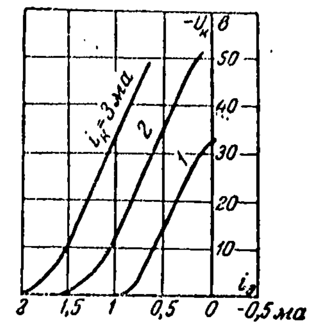

Fig. 17-16. Static amplification characteristics of a point-contact triode.

output power, limited operating frequencies, and also their low operating temperature, a change in which considerably changes their parameters.

Studies of the first samples of semiconductor triodes showed that the amplification factor is considerably decreased by a rise in frequency of from 1 to 10 Mc. At the present time, it is possible to make semiconductor triodes which can operate on frequencies 150-300 Mc. This limit is conditional upon the time required for electrons or holes to traverse the distance between the emitter and the collector, resembling the principle that the frequency limit of vacuum triodes is conditional upon the transit time of the electrons between the cathode and the grid.

Among the characteristics of point-contact triodes, the static characteristics and especially the output characteristics and amplifi-

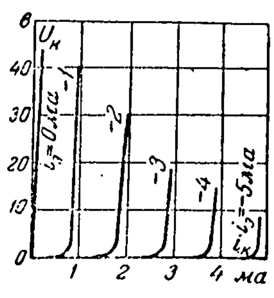

Fig. 17-17. Static output characteristics of an N-P-N junction triode for high current and voltage values.

cation characteristics, have practical value. The output characteristics are a factor of the dependence of negative voltage at the collector on the collector's negative current. A family of these characteristics is shown in Fig. 17-15. One may see that the output characteristics of the point-contact triode are extremely similar to the characteristics of a vacuum triode.

The amplification characteristics express the dependence of negative voltage at the collector on emitter current (Fig. 17-16). They resemble the grid characteristics of a vacuum triode. The basic difference of the characteristics of a semiconductor triode from the characteristics of a vacuum triode consists of the fact that voltages are plotted along the Y axis and currents along the X axis. Therefore, the quantities determined from these characteristics and characterizing the properties of semiconductor triodes have the dimensions of resistances.

It is possible to compute resistances r_{22} and r_{21} from the resistances for a selected condition.

In practice, it has also turned out to be convenient to construct the characteristics of junction triodes, which are not essentially different from the characteristics of a point-contact triode, on the same coordinates. The static output characteristics of a junction triode for large currents and voltages are shown in Fig. 17-17. If these characteristics are constructed by plotting current on the Y axis and voltage along the X axis, that is, as if for a vacuum tube, we obtain characteristics resembling the characteristics of a pentode. In general, the method of constructing the characteristics is determined by the same formulas used to calculate one or the other of the parameters.

17-6. PHOTOVARISTORS.

A semiconductor device which changes its electrical resistance

under the action of light is called a photovaristor.

In photovaristors, when photon energy is absorbed by a semiconductor, the electrons do not leave the semiconductor as will occur during extrinsic photoeffect, but traverse only from the space-charge energy region into the conduction band; holes are also formed in the space-charge region. The occurrence of electrons in the conduction band and holes in the space-charge region creates a sharp reduction of the electrical resistance of the semiconductor during its illumination (intrinsic photoeffect). However, one should bear in mind that the electrical resistance of a semiconductor may be decreased only if the photon energy $h\nu$ is equal to or greater than the energy ΔE_0, which electrons must have to overcome the forbidden zone, that is, $h\nu \geqslant \Delta E_0$. If $h\nu < \Delta E_0$, the electrons cannot overcome the forbidden zone and traverse from the space-charge region to the conduction band and photoconduction does not occur.

Thus, a radiation frequency ν_0 exists for every semiconductor, in the presence of which and above which photoconduction takes place. If the frequency of the radiation falling on the semiconductor is less than frequency ν_0, photoconduction does not occur. Therefore, it follows that photoconduction is determined both by the quantity of energy ΔE_0 characterizing the semiconductor and by the quantity of ν, characterizing the radiation.

Photovaristors, depending upon their purpose, are made of various semiconductor materials: selenium, thallium sulphide, lead sulphide, and others.

A simplified construction of a selenium photovaristor is shown in Fig. 17-18 as an example. A system of closely arranged parallel grooves (or a system of electrodes in the form of strips) is engraved on the surface of a glass plate. The distance between the grooves is about

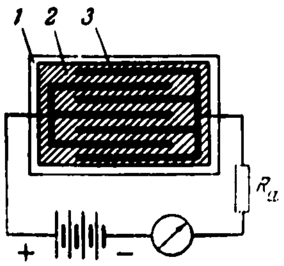

Fig. 17-18. Simplified diagram of structure and circuitry of a selenium photovaristor. 1) Glass plate; selenium; 3) conductive grooves.

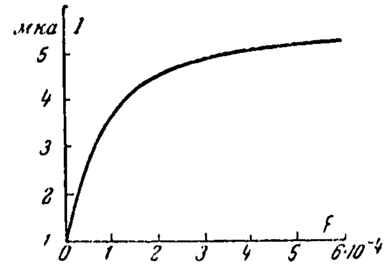

Fig. 17-19. Light characteristics of a thallium-sulphide photovaristor. 1) μa.

0.1 mm. Then the grooves are filled with a conductor substance (for example, graphite or gold) and a thin layer of a semiconductor is laid on the surface of the plate. If the conductive grooves are joined one to the other and two terminals are led from them, a system of closely arranged parallel conductors is obtained on the glass plate, the intervals between which are covered with a semiconductor. These intervals are also photovaristors. Naturally, the shorter the distance between the grooves, the smaller the semiconductor's resistance.

Since the semiconductors used as photovaristors strongly change their properties in contact with the air, photovaristors are placed in a glass bulb which has been either evacuated or filled with an inert gas for more stable operation.

The dark resistance of a selenium photovaristor ordinarily amounts to 0.2-0.5 megohm. Hence, the dark current is small. When the photovaristor is illuminated, the dark resistance is decreased by 3 to 5 times and the current increases correspondingly. The dark resistance of thallium photovaristors amounts to several megohms. Therefore, their dark current is considerably less, by comparison, than selenium photovaristors.

The basic characteristics of photovaristors are the light, spectral, and frequency characteristics.

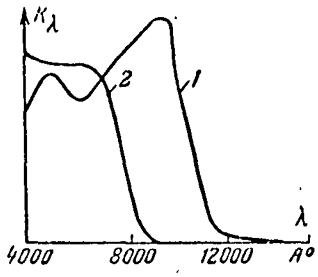

Fig. 17-20. Spectral characteristics of a (1) thallium-sulphide and a (2) selenium photovaristor.

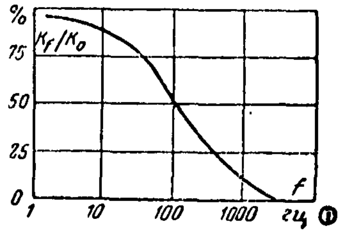

Fig. 17-21. Frequency characteristics of a selenium photovaristor.
1) cps.

The light characteristics of a thallium-sulphide photovaristor are nonlinear (Fig. 17-19). When the luminous flux is increased, at first current increases relatively fast, but later ever slower and slower. This dependence of current upon luminous flux is roughly expressed by the formula $I \approx k\sqrt{F}$.

The spectral characteristics of selenium and thallium-sulphide photovaristors are shown in Fig. 17-20. They show that the selenium photovaristor is most sensitive to rays of the visible and red part of the spectrum, and the thallium-sulphide photovaristor to rays of the infrared part of the spectrum.

The frequency characteristics of a selenium photovaristor, expressing the dependence of sensitivity of the photovaristor on the frequency of change of the luminous flux, are shown in Fig. 17-21. In a given case, the sensitivity of the photovaristor falls with the increase of frequency change of the luminous flux, that is, the quantity of photocurrent does not instantly follow the change of the luminous flux, but lags behind more and more as frequency change of the luminous flux increases, which is a great shortcoming of photovaristors and limits their field of usage to the limits of the smaller frequency changes of the luminous flux. The photovaristors are not used with rapid luminous flux changes.

The sensitivity of a selenium photovaristor, in the steepest part

of its light characteristics corresponding to a small luminous flux, attains values of 800 to 1000 $\mu a/\underline{l}$, but with increase of luminous flux it decreases the value to 150 to 200 $\mu a/\underline{l}$.

Thallium-sulphide (thallofide) photovaristors are made of thallium sulphide. In manufacturing photovaristors, thallium sulphide is oxidized at 80°C in the air for a long time. Thallium-sulphide photovaristors are supplied with special red filters which protect the light-sensitive layer from the destructive action of short-wave light.

The sensitivity of thallium-sulphide photovaristors at very small luminous fluxes reaches 10,000 $\mu a/\underline{l}$.

17-7. BARRIER-LAYER PHOTOCELLS.

A semiconductor device in which the action of light originates emf, which creates electrical current in the external circuit, is called a barrier-layer photocell (or a barrier photocell).

In contrast to a photovaristor, which consists only of one semiconductor, a barrier photocell consists of two semiconductors, one of which has electron conduction and the other hole conduction. Both semiconductors are divided by the so-called P-N junction, which transmits current only in one direction, that is, performs the function of a barrier layer. Selenium, thallium sulphide, cuprous oxide, silver sulphide, and other semiconductor materials are used in manufacturing barrier photocells. Therefore, depending upon the type of material used, there are selenium, thallium-sulphide, silver-sulphide, and other types of photocells.

Depending upon temperature treatment, the P-N junction (the barrier layer) may be formed either between the upper semitransparent metallic layer, through which light penetrates into the semiconductor, and the semiconductor, or between the lower metallic plate, which serves as a

base, and the semiconductor layer. Therefore, there are two types of barrier-layer photocells: photocells with front-wall (front) photoeffect, in which the barrier layer is formed between the semiconductor and the semitransparent metallic layer, and photocells with rear-wall (rear) photoeffect, in which the barrier layer is formed between the semiconductor and the metallic base plate. Cuprous-oxide photocells have front and rear photoeffect; a simplified diagram of the structure of such photocells is shown in Fig. 17-22.

Let us consider the operating principle of a barrier photocell.

In Section 17-2, we discussed the fact that some internal difference of potential (a potential threshold) is established between the electron and hole regions of a semiconductor (in the P-N junction) as a result of the diffusion of major carriers of charges in mutually opposite directions (Fig. 17-5a). By this means, the internal electrical field E in the P-N junction is moved from the electron region to the hole region. If the semiconductor is subjected to the influence of light, electrons and holes are originated in it in equal amounts: the electrons, having absorbed the photons' energy, transit from the space-charge energy region to the conduction band; as a result of the movement of the electrons, holes are formed in the space-charge region, in a quantity equal to the number of electrons which have transited into the conduction band.

Under the action of electrical field E, which was created by the division of the majority carriers of charges, the minority carriers, electrons in the hole region and holes in the electron region, which were also created by radiation action, are accelerated in the direction of the P-N junction and are shifted across it in mutually opposite directions: the holes to the hole region and electrons to the electron region. As a result of this, the hole region's potential is increased,

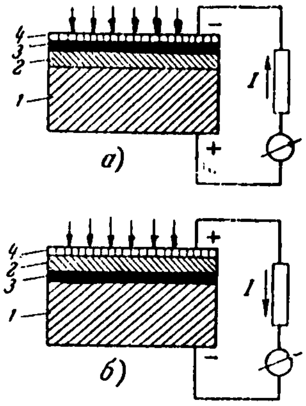

Fig. 17-22. Structure of a cuprous-oxide barrier-layer photocell.
1) Copper; 2) cuprous oxide; 3) barrier layer; 4) semitransparent layer of metal.

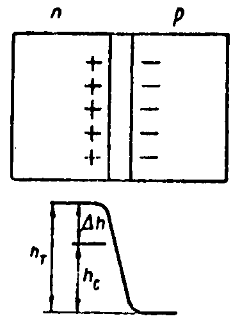

Fig. 17-23. Change of potential threshold by creation of photoelectric motive force.

and the electron region's potential is decreased. Simultaneously, an electrical field, conditioned by the division of the minority carriers of charges which were formed by the radiation action, is originated in the P-N junction. This field blocks further division of the above-mentioned minority carriers, and, consequently, further change of the potentials of the electron and hole regions. At the end, a state of equilibrium sets in, in which the number of charges traversing in one direction in a given unit of time is equal to the number of charges traversing in the opposite direction. As a result, the height of the potential threshold h_t,* which corresponded to the unilluminated semiconductor, is decreased to value h_c,** corresponding to the illuminated semiconductor (Fig. 17-23). Thus, the influence of light on the semiconductor on one hand leads to the establishment of a state of equilibrium of the minority carriers and on the other hand to disruption of the state of equilibrium of the majority carriers, as a result of

* [h_t - h_t - $h_{temnovoe}$ - h_{dark}.]
** [h_c - h_s - $h_{svetovoe}$ - h_{light}.]

which the height of the potential threshhold is decreased by quantity $h_t - h_c = \Delta h$.

The greater the light intensity, the greater the number of electrons and holes that originate in the semiconductor and the greater the difference Δh is increased. This difference, originated under the action of light in a barrier photocell, has been named photoelectric motive force.

When illumination of the semiconductor ceases, height of potential threshold h_c again receives the value h_t, by means of which difference $h_t - h_c$ and, consequently, photoelectric motive force drops to zero.

Thus, the quantity of photoelectric motive force changes as luminous flux is changed. With minor changes of luminous flux this change bears a linear character. With greater luminous fluxes photoelectric motive force is increased more slowly than luminous flux as a consequence of which linearity is disrupted. With sufficiently great values of luminous flux, saturation sets in: photoelectric motive force is barely increased by an increase in luminous flux. This is explained by the fact that with great luminous fluxes the number of charges created in the semiconductor by the action of light attain limit value, and are increased very little by further increase of luminous flux.

In the presence of photoelectric motive force, the electrical field in the P-N junction is moved, as is the E field, from the electron conduction region of the semiconductor to the region of the semiconductor with hole conduction. Therefore, an electrode in contact with the hole semiconductor always has a positive potential, and an electrode in contact with an electron semiconductor, negative potential. From this it follows that in the same types of barrier photocells, the semitransparent electrode may have either positive or negative potential depending upon which semiconductor it contacts.

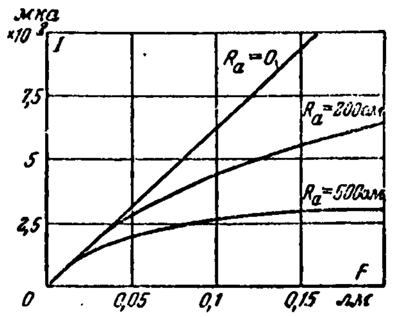

Fig. 17-24. Light characteristics of a thallium-sulphide barrier-layer photocell at different values of load resistance.

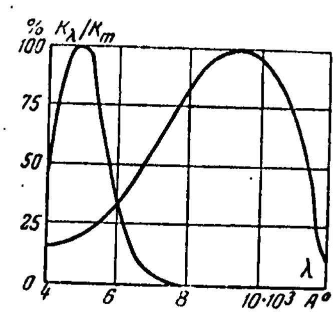

Fig. 17-25. Spectral characteristics of thallium-sulphide (1) and cuprous-oxide (2) barrier-layer photocell.

Thus, barrier photocells differ from photocells with extrinsic and intrinsic photoeffect by the presence of an intrinsic emf.

The basic characteristics of barrier photocells are light and spectral characteristics. The light characteristics of a thallium-sulphide photocell, obtained at various values of load resistance R_a, are shown in Fig. 17-24. These characteristics show that when $R_a = 0$ the photocurrent changes linearly with change of luminous flux, but when $R_a > 0$ the relationship $I = f\ (F)$ becomes nonlinear. The light characteristics of a selenium photocell also have an analagous form.

The spectral characteristics of thallium-sulphide and cuprous-oxide barrier-layer photocells are shown in Fig. 17-25. The spectral characteristics of a thallium-sulphide photocell includes the region of the visible spectrum and the greater part of the infrared rays (the maximum sensitivity corresponds to wavelengths of the order of 9,500 A), and cuprous-oxide photocells are sensitive to blue rays. The sensitivity of cuprous-oxide photocells with frontal photoeffect amounts to 100-200 μa/l, and of selenium photocells to 400-500 μa/l with the maximum in the regions of the blue and the green rays.

In the rear photoeffect, when the semitransparent film is positively charged and the basic plate negatively, thallium-sulphide photocells

have a very great sensitivity (from 4,000 to 6,000 μa/1).

17-8. APPLICATIONS OF SEMICONDUCTOR DEVICES.

Germanium junction diodes are chiefly used as rectifiers. The low value of rectified voltage is a substantial shortcoming of these diodes. However, a series connection of diodes having equal internal resistances both for forward and reverse current permit the attainment of a rectified

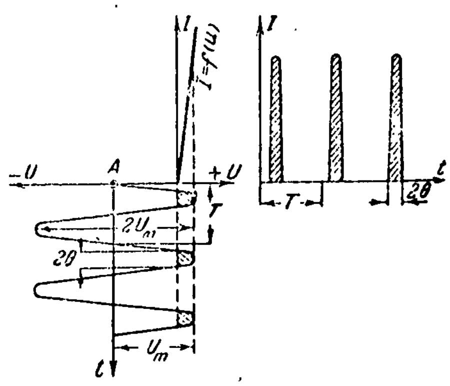

Fig. 17-26. Detection with a small angle of cutoff θ, without considering reverse conduction, and aligning the working section of the characteristics I = f(u).

voltage of the order of 200-250 v, which is entirely adequate for supply of tube anode circuits.

Germanium diodes are also used in rectifying circuits intended for supplying low-power receiving and measuring units. Diodes with high breakdown voltage and insignificant resistance in a forward direction are used in such circuits.

Point-contact diodes, both silicon and germanium, are used primarily as detectors and mixers in receivers in the microwave band. The germanium diode is used for detection of AM videosignals in television sets.

In detection, semiconductor diodes often operate with small angles of cutoff θ, when the quiescent point A is biased to the left (Fig.

17-26) and the quantity of this bias is almost equal to the amplitude U_t* of the oscillations detected. In such a condition the voltage led to the diode is almost equal to the doubled signal amplitude. Therefore, to detect the oscillations of a signal with an amplitude, for example, of 25 v, a diode which can stand a voltage of not less than 50 v is required.

Diodes of the DG-Ts type may also be used for generation of oscillations with a power of the order of 100-120 mw at frequencies of from 1 to 2 kc. At higher frequencies, oscillation power is sharply reduced because of heat losses in the semiconductor. Besides this, these diodes are used in circuits for doubling frequencies of low power. Finally, germanium diodes of the DG-Ts type are widely utilized in measuring techniques when measuring high frequency voltages and currents (up to 100 Mc).

Semiconductor rectifiers are not limited to two types: diodes and triodes. At the present time, experimental forms of semiconductor tetrodes have been developed.

Many circuits for generators, low- and medium-frequency amplifiers, direct-current amplifiers, videosignal amplifiers in television engineering, and others, are accomplished by means of semiconductor rectifiers and amplifiers. The perfection of printed-circuit technique and the application of crystal diodes, triodes, and tetrodes in them create the possibility of considerable decrease in power requirements, external dimensions, and weight of such equipment.

Semiconductor devices have found the most diversified application in radio engineering, automation, telemechanics, and in many other fields of science and technology. Even now, when they are, in essence, in the initial stages of their development, it is possible to say that they have a great future, since their role in science and technology is

* [U_T - U_t - U_{tok} - $U_{current}$.]

steadily growing. In some fields these devices are successfully replacing the ordinary vacuum tube, and in other fields they are also successfully being used together with vacuum tubes. Naturally, we cannot say at present that semiconductor rectifiers and amplifiers will "entirely supersede vacuum tubes." In fact, semiconductor diodes, triodes, and tetrodes are supplementing vacuum tubes very well, and, together with the latter, permit new solutions of various scientific and engineering problems.

Photovaristors and barrier-layer photocells are mainly used only in visible light engineering measuring equipment, in visual light signaling apparatus, and in other equipment which operates by slow changing of luminous flux and in some cases without preamplification of the photocurrent, since the current of these devices is adequately high. The limitation of the field of use of photovaristors and barrier-layer photocells is explained by the fact that they have great inertness in operation with rapidly-changing luminous fluxes. Besides this, the field of application of photovaristors and barrier-layer photocells is also limited by reason of the strong dependency of these devices' current and emf on the temperature.